

P O L S K A A K A D E M I A N A U K
I N S T Y T U T F I Z Y K I

ACTA PHYSICA POLONICA

DWUMIESIĘCZNIK

Vol. XVI — Fasc. 3

WARSZAWA 1957

Orders and inquires concerning
Acta Physica Polonica
— complete sets, volumes and single fascicules —
as well as other

Polish scientific periodicals
published
before and after the war,
regularly and irregularly,
are to be sent to:

Export and Import Enterprise „RUCH”
Warszawa 1, P.O. Box 154, Poland
Ask for catalogues, folders and sample copies.

THE EQUATIONS OF MOTION IN GENERAL RELATIVITY THEORY AND THE ACTION PRINCIPLE

BY L. INFELD

(Received December 17, 1956)

This paper deals with the formulation of the equations of motion for p bodies, assuming: 1. the field equations of General Relativity Theory, 2. the fact that the energy-momentum tensor is different from zero only along certain world lines; this means that this tensor depends linearly on Dirac's δ -functions.

It can be shown that from these simple assumptions the proper Lagrangian can be found almost without tedious calculation. It will give the equations of motion up to the post-Newtonian order.

Also the connection between these equations and those of the "geodetic lines" is elucidated.

§ 1 Historical introduction

The history of the problem of motion in general relativity theory usually, and rightly so, starts with the paper written by Einstein and Grommer in the year 1927. This paper showed for the first time that the equations of motion for a testing particle (with mass $m \rightarrow 0$) need not be added to the field equations — that they can be deduced from the relativistic field equations. For many years afterward, Einstein, and then Einstein with his collaborators, tackled the problem of the motion of two particles; to find whether their equations of motion can be deduced from the field equations. The answer to this question was given in the paper by Einstein, Infeld and Hoffmann in 1938 in which the two-body problem was solved for the first time.

Independently of us and a little later Fock (1939) also deduced the equations of motion, though only the Newtonian ones, from the field equations. Later Papapetrou (1951) simplified his procedure and deduced the post-Newtonian equations of motion, after Petrova (1949) had done the same thing on the basis of Fock's theory. Petrova's and Papapetrou's results were the same as ours.

What are the essential similarities and differences between Einstein's theory, especially as formulated in the two later papers (Einstein-Infeld 1939 and 1949) and the Fock-Papapetrou papers?

Without going into detail we can say in general that they have one essential idea in common but two different ideas, of which I regard only one as essential.

The idea common to both sets of papers is the approximation method about which I shall say more later.

Now, about the first difference:

In the school represented by Einstein we take as our field equations those of empty space, that is¹ in the usual notation:

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = 0. \quad (1.1)$$

Einstein always thought that to use

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = -8\pi T_{\alpha\beta} \quad (1.2)$$

($T_{\alpha\beta}$ being the energy-momentum tensor) instead of (1.1) is somehow in bad taste; because we do not know in (1.2) what $T_{\alpha\beta}$ is and we mix a geometrical tensor on the left side with a physical tensor on the right side. This was the reason for Einstein's long search for a unified field theory in which such a mixture of physics and geometry would not appear.

We know that there is no solution of (1.1) representing spread out continuous matter. Therefore by assuming (1.1) we represent matter by means of singularities. The method first used by us, consisted in forming certain two-dimensional surface integrals over surfaces enclosing these singularities. The field equations prescribed the laws by which the surfaces enclosing the singularities and hence these singularities moved. Therefore these laws were — until the post-Newtonian approximation — deduced from the field equations.

However Fock and Papapetrou consider (1.2) and use definite expressions for $T_{\alpha\beta}$.

This difference does not seem to me to be an essential one for the following reasons: we do not know the real distribution of matter. Neither of these methods depicts reality properly. The use of our method based on (1.1) means: if the two bodies are a great distance apart so that we may assume, approximately, central symmetry of the field near one body, then the exact knowledge of density distribution inside the enclosing surface is not essential. Outside the enclosing surface, (1.1) is valid.

I would like to characterise this difference between the two schools by an example of a simple situation: that is classical gravitational theory. There we have two kinds of equations depending on whether matter is represented by singularities or by a continuous distribution. In the first case we have Laplace's equation

$$\Delta\varphi = 0, \quad (1.3)$$

¹ Greek indices run from 0 to 3. Latin from 1 to 3. Repetition implies summation. The quadratic form for a geodetic coordinate-system is

$$ds^2 = \eta_{\alpha\beta} dx^\alpha dx^\beta; \quad \eta_{00} = 1; \quad \eta_{0m} = 0; \quad \eta_{mn} = -\delta_{mn}.$$

The velocity of light $c = 1$.

in the second case Poisson's equation

$$\Delta\varphi = 4\pi\rho. \quad (1.4)$$

Now it is more common to write Laplace's equation, when thinking about its spherically symmetric in the form

$$\Delta\varphi = 4\pi\delta_{(3)} \quad (1.5)$$

where m is the mass and $\delta_{(3)}$ is the three-dimensional Dirac δ — function.

Let us try to use the δ -functions consistently in general relativity theory. Here, too, we have, in the first approximation, using (1.1) for empty space, the Laplace equation which we solve by a central symmetric solution. This means that a solution of (1.1), in the first approximation, valid everywhere, corresponds to that of (1.2) where the energy momentum tensor $T_{\alpha\beta}$ is proportional to Dirac's δ -function. Thus the use of (1.2) with $T_{\alpha\beta}$ proportional to $\delta_{(3)}$, corresponds exactly to our previous considerations of (1.1) with singular solutions; just as equation (1.5) is only a different form of (1.3).

Indeed the use of (1.2) with $T_{\alpha\beta}$ proportional to $\delta_{(3)}$ -functions tremendously simplifies the entire deduction of the equations of motion. This simplification was achieved in my paper of 1954, but the entire procedure can still be clarified and simplified. In fact I shall present here the derivation of the post-Newtonian equations of motion with almost no tedious calculations².

Summarizing, I may say, that though I believe in Einstein's idea of not using the energy momentum tensor, yet I was unfaithful to it, because I used the energy-momentum tensor as proportional to the $\delta_{(3)}$ -function, to express the singular solutions.

The next difference between Fock's and Einstein's school is more essential.

Fock, Petrova and Papapetrou use the harmonic coordinate system, that is the four equations:

$$\left[(-g)^{\frac{1}{2}} g^{\alpha\beta}\right]_{,\beta} = \frac{\partial(-g)^{\frac{1}{2}} g^{\alpha\beta}}{\partial x^{\beta}}, \quad (1.6)$$

As a matter of fact, Fock considers the choice of this coordinate system to be extremely important, claiming that its addition to the gravitational equations (plus some conditions at infinity) restricts the coordinate system up to a Lorentz transformation. Thus, for Fock, the choice of the harmonic coordinate condition becomes a fundamental law of nature changing the character of Einstein's general relativity theory into a theory of the gravitational field, valid only in inertial coordinate systems. Other (like Papapetrou 1951) who based their research on Fock's work, do not go so far, but they also regard the coordinate condition (1.6) as essential for the deduction of the equations of motion.

² Some of the ideas presented here are due to Mr. Plebański and myself; they will appear in fuller form in a book "Relativity and Motion" which we are writing.

In our first paper (Einstein, Infeld and Hoffmann 1938) we used a coordinate condition different from the harmonic one; yet we obtained the same equations of motion later obtained by Papapetrou and Petrova. In a few subsequent papers (Einstein and Infeld 1949, Infeld 1953, 1954, Teisseyre 1954) we carefully analysed the problem.

I maintain that the equations of motion have nothing whatever to do with harmonic coordinate conditions; they have much to do with the method of approximation. Indeed it is this method which uniquely determines the equations of motion up to the post-Newtonian order. I will show explicitly later that the violation of the harmonic coordinate condition does not change the post-Newtonian equations of motion as long as we stick to the approximation procedure.

§ 2. Some notations and mathematical preliminaries

We have a world line $\xi^k(t)$ and a field, say a scalar field φ that depends on coordinates x^k and on time $x^0 = t$ and also on the $\xi^k(t)$ and their time derivatives:

$$\varphi = \varphi(x^k, t, \xi^k, \dot{\xi}^k); \quad \dot{\xi}^k = \frac{d\xi^k}{dt}. \quad (2.1)$$

We assume that near the line $\xi^k(t)$ the field becomes singular and has the following form:

$$\varphi = \frac{-1}{\varrho} + \tilde{\varphi} + \tilde{\varphi}_{,s} (x^s - \xi^s) + \frac{1}{2} \tilde{\varphi}_{,sr} (x^s - \xi^s) (x^r - \xi^r) + \dots \quad (2.2)$$

Here

$$\varrho^2 = (x^s - \xi^s) (x^s - \xi^s) \quad (2.3)$$

and, therefore

$$\tilde{\varphi} = \left(\varphi - \frac{-1}{\varrho} \right) \Big|_{x^s = \xi^s}. \quad (2.4)$$

Similarly

$$\tilde{\varphi}_{,s} = \frac{\partial}{\partial x^s} \left(\varphi - \frac{-1}{\varrho} \right) \Big|_{x^s = \xi^s}. \quad (2.5)$$

We must distinguish between

$$\tilde{\varphi}_{,s} \text{ and } \tilde{\varphi}_{,s} = \frac{\partial \tilde{\varphi}}{\partial \xi^s} = \tilde{\varphi}_{, \xi^s} \quad (2.6)$$

which generally are not equal to each other.

If φ were not singular in the curve $\xi^k(t)$ we could have defined φ in the following way:

$$\tilde{\varphi} = \int \varphi \delta_{(3)}(x^s - \xi^s) d_{(3)}x. \quad (2.7)$$

Now we can narrow the definition of Dirac's δ -functions in such a way that (2.7) remains true even if φ has a singularity up to the k 'th order. Indeed we can show (see Appendix A) that such δ -functions can be constructed as limits of ordinary functions. This is convenient because, by the use of such δ -functions we can get rid of infinities without recourse to the renormalisation procedure. Thus all the δ -functions used here will have the property of changing φ into $\tilde{\varphi}$, where $\tilde{\varphi}$ is a continuous function of the $\xi^k, \dot{\xi}^k, \ddot{\xi}^k$.

We can, therefore, use (2.7) as the definition of $\tilde{\varphi}$, where δ is the three-dimensional Dirac's δ -function, satisfying the following conditions:

1) $\delta(x)$ can be treated formally as a spherically symmetric function for which all the derivatives exist.

2) $\delta(x) = 0$ for $x \neq 0$.

3) For every continuous $f(x)$ in the arbitrary region $\Omega(\tilde{x}_0)$ forming a neighbourhood of \tilde{x}_0 , we have:

$$\int_{\Omega(\tilde{x}_0)} d_{(3)}x \delta(\tilde{x} - \tilde{x}_0) f(\tilde{x}) = f(\tilde{x}_0).$$

4) For an arbitrary neighbourhood $\Omega(o)$ of the point $\tilde{x}_0 = 0$, we have

$$\int_{\Omega(o)} d_{(3)}x \delta(\tilde{x}) |\tilde{x}|^{-p} = 0, \text{ for } p = 1, 2, \dots, k.$$

It is the fourth condition that distinguishes *this* δ -function from the usual Dirac's δ -function. The proof of the existence of such functions is given in Appendix A. (See also Infeld and Plebański 1956 and 1957).

Thus the "tweedling" means two things: firstly, the singularities are ignored; secondly, for x^k the ξ^k 's are introduced.

We can now return to the problem of the difference between $\tilde{\varphi}_{,s} = \frac{\partial \tilde{\varphi}}{\partial \xi^s}$ and $\tilde{\varphi}_{,s}$. We have from (2.6)

$$\boxed{\tilde{\varphi}_{,\xi^s} = \tilde{\varphi}_{,\xi^s} + \tilde{\varphi}_{,s}} \quad (2.8)$$

This means: $\tilde{\varphi}_{,s}$ and $\tilde{\varphi}_{,\xi^s}$ are equal if and only if $\tilde{\varphi}_{,\xi^s} = 0$. This is certainly so if the part of φ that gives a contribution to φ does not depend on ξ^s .

One more formula that will play an important role later and which follows from the definition (2.7) is:

$$\boxed{\tilde{\varphi}_{,0} = \frac{\partial \tilde{\varphi}}{\partial t} = \tilde{\varphi}_{,0} + \tilde{\varphi}_{,s} \dot{\xi}^s = \tilde{\varphi}_{,\alpha} \dot{\xi}^\alpha} \quad (2.9)$$

Let us now assume two functions e.g. φ and ψ :

$$\begin{aligned}\varphi &= \frac{\varphi}{\varrho} + \tilde{\varphi} + \tilde{\varphi}_{,s} (x^s - \xi^s) + \frac{1}{2} \tilde{\varphi}_{,sr} (x^s - \xi^s) (x^r - \xi^r) + \dots \\ \psi &= \frac{\psi}{\varrho^3} + \frac{\psi}{\varrho} + \tilde{\psi} + \tilde{\psi}_{,s} (x^s - \xi^s) + \dots\end{aligned}\quad (2.10)$$

We have in this case:

$$\tilde{\varphi}\tilde{\psi} = \tilde{\varphi}\tilde{\psi} \quad (2.11)$$

Indeed forming $\varphi\psi$ and ignoring singular expressions and those that become zero for $x^s = \xi^s$, we are left with the following expressions of the order zero in $(x^s - \xi^s)$:

$$\tilde{\varphi}\tilde{\psi} + \tilde{\varphi}_{,s} \varphi^{-1} \frac{x^s - \xi^s}{\varrho} + \tilde{\psi}_{,s} \psi^{-1} \frac{x^s - \xi^s}{\varrho} + \frac{1}{6} \tilde{\varphi}_{,srp} \psi^{-3} \frac{(x^s - \xi^s)(x^r - \xi^r)(x^p - \xi^p)}{\varrho^3}.$$

The „Tweedling” of these expressions gives us:

$$\begin{aligned}\tilde{\varphi}\tilde{\psi} &= \int \delta_{(3)} d_{(3)}x \left[\tilde{\varphi}\tilde{\psi} + (\tilde{\varphi}_{,s} \psi^{-1} + \tilde{\psi}_{,s} \varphi^{-1}) \frac{x^s - \xi^s}{\varrho} + \right. \\ &\quad \left. + \frac{1}{6} \tilde{\varphi}_{,srp} \psi^{-3} \frac{(x^s - \xi^s)(x^r - \xi^r)(x^p - \xi^p)}{\varrho^3} \right].\end{aligned}$$

The last two expressions under the integral sign give zero, because they are product of *symmetric* δ -functions and *odd* powers of $(x^s - \xi^s)$. Thus we obtain:

$$\tilde{\varphi}\tilde{\psi} = \int \tilde{\varphi}\tilde{\psi} \delta_{(3)} d_{(3)}x = \tilde{\varphi}\tilde{\psi}.$$

But (2.11) would not be true if φ or ψ had a singularity of order ϱ^{-2} . Thus (2.11) is true if ϱ appears to an *odd* power in the singular parts of φ and ψ . Therefore we have to be cautious in applying the last equation.

Thus we shall have vectors and tensors defined only *along* the curve, like \tilde{S}^a , $\tilde{T}^{a\beta}$ etc. We can define the metric tensor *along* the curve $\tilde{g}_{a\beta}$ and $\tilde{g}^{a\beta}$ and assuming (2.11) we have

$$\tilde{g}_{a\alpha} \tilde{g}^{\beta\alpha} = \tilde{g}_{a\alpha} \tilde{g}^{\beta\alpha} = \delta_a^\beta \quad (2.12)$$

To such tensors we can apply tensor algebra and tensor analysis but *only* along the curve.

Since

$$\tilde{g}_{a\beta} = \int_{\Omega_{(3)}} g_{a\beta} \delta_{(3)} d_{(3)}x \quad (2.13)$$

(where $\Omega_{(3)}$ is a small three-dimensional neighbourhood surrounding the singularity), $\tilde{g}_{a\beta}$ will be a tensor along the curve if $d_{(3)} d_{(3)}x$ is an invariant. Thus we must discuss the transformation properties of $\delta_{(3)}$.

The four-dimensional relativistic Dirac's $\delta_{(4)}$ is a scalar density because of the invariant equation

$$\int \delta_{(4)} d_{(4)}x = 1. \quad (2.14)$$

Now let us take

$$\delta_{(4)} = \delta_{(4)}(x^a - \xi^a) \quad (2.15)$$

where

$$\xi^a = \xi^a(\lambda) \quad (2.16)$$

λ being an invariant parameter. Then we can form an invariant density function

$$\int_{-\infty}^{+\infty} \delta_{(4)} d\lambda = \int_{-\infty}^{+\infty} \delta_{(4)} \frac{d\lambda}{d\xi^0} d\xi^0 = \frac{d\lambda}{dt} \delta_{(3)}. \quad (2.17)$$

As far the transformation properties are concerned this is the definition of $\delta_{(3)}$. Thus $\delta_{(3)}$ is the zero component of a density vector. From this definition it follows that:

$$\int_{\Omega_{(3)}} \delta_{(3)} d_{(3)}x = \int_{\Omega_{(3)}} \delta_{(3)} \frac{d\lambda}{dt} d_{(3)}x \frac{dt}{d\lambda} = \int_{\Omega_{(3)}} \int_{-\infty}^{+\infty} \delta_{(4)} d_{(3)}x dt = \int_{\Omega_{(4)}} \delta_{(4)} d_{(4)}x. \quad (2.18)$$

Thus the space integral of $\delta_{(3)}$ is an invariant.

One more remark concerning the notation: If we have many curves we shall distinguish between them by the index written above: ξ^a ; $a = 1, 2, \dots, p$. Then we should also write $\tilde{\varphi}^a$, meaning:

$$\tilde{\varphi}^a = \int_{\Omega_{(3)}} \varphi^a d_{(3)}x; \quad \delta^a = \delta_{(3)}(x^s - \xi^s) \quad (2.19)$$

However, for the sake of simplicity we shall not write a above the "tweedle" always understanding that it means the *first* (or only) curve: $\tilde{\varphi} = \tilde{\varphi}^1$.

§ 3. The gravitational equations

The gravitational equations expressed in contravariant tensor densities are:

$$G^{a\beta} = R^{a\beta} - \frac{1}{2} g^{a\beta} R = -8\pi T^{a\beta}. \quad (3.1)$$

In our case of, say, two particles moving along lines ξ^s and ξ^s we have:

$$T^{a\beta} = t^{a\beta} \delta^1 + t^{a\beta} \delta^2. \quad (3.2)$$

Before going further we shall introduce an invariant and finite line element, concerning the first particle (without writing the "one" above).

$$\widetilde{ds}^2 = \widetilde{g}_{\alpha\beta} d\xi^\alpha d\xi^\beta. \quad (3.3)$$

From the tensor density we can now form a *tensor* along a curve (the first one)

$$\frac{dt}{\widetilde{ds}} \int_{\Omega(x)} T^{\alpha\beta} d_{(3)}x = \widetilde{t}^{\alpha\beta} \frac{dt}{\widetilde{ds}}. \quad (3.4)$$

What is $t^{\alpha\beta}$? We shall see in the next section that the answer to this question is a unique one; that the consistency of the field equations demands the following equations to be true (Tulczyjew 1957):

$$t^{\alpha\beta} \frac{dt}{\widetilde{ds}} = \mu \xi^{\alpha'} \xi^{\beta'}; \quad \xi^{\alpha'} = \frac{d}{\widetilde{ds}} \xi^{\alpha}; \quad \mu = \text{rest mass} \quad (3.5)$$

We shall show (§ 4) that not only has $t^{\alpha\beta}$ the form (3.5) but also that the rest mass μ is constant. For the moment, however, we shall use (3.5) as an *assumption*, without stipulating that μ in (3.5) is a constant.

From (3.5) follows:

$$t^{\alpha\beta} = \mu \dot{\xi}^{\alpha} \dot{\xi}^{\beta} = m \dot{\xi}^{\alpha} \dot{\xi}^{\beta}; \quad \dot{\xi}^{\alpha} = \frac{d}{dt} \xi^{\alpha} \quad (3.6)$$

$$\boxed{\frac{m}{\mu} = \frac{dt}{\widetilde{ds}}}$$

and

$$\boxed{T^{\alpha\beta} = \sum_{a=1}^p m^a \xi^{\alpha a} \dot{\xi}^{\beta a} \delta^a} \quad (3.7)$$

The right hand side of the gravitational equations is uniquely determined by the condition that $T^{\alpha\beta}$ depends linearly on the δ^a 's.

§ 4. The general equations of motion

As a consequence of Bianchi's identities we always have

$$G^{\alpha\beta}_{;\beta} = \left(R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R \right)_{;\beta} = -8\pi T^{\alpha\beta}_{;\beta} = 0 \quad (4.1)$$

where the semicolon indicates covariant differentiation. Taking the integral of $T^{\alpha\beta}_{;\beta}$ over the three-dimensional region surrounding the first singularity and multiplying

by $\frac{dt}{ds}$, we have:

$$\frac{dt}{ds} \int_{\Omega_{(3)}} T^{\alpha\beta}_{;\beta} d_{(3)}x = \tilde{A}^\alpha = 0, \quad (4.2)$$

where \tilde{A}^α is a vector defined through (4.2) along the *first* curve. Generally we shall have

$$\tilde{A}^a = 0; \quad \alpha = 0, 1, 2, 3; \quad a = 1, 2, \dots, p \quad (4.3)$$

if p is the number of singularities. Thus in (4.3) we have as a consequence of the field equations, $4p$ equations which we shall call the *equations of motion* of the p singularities. Indeed they contain $4p$ unknowns:

$$\xi^a(t) \text{ and } m(t); \quad s = 1, 2, 3; \quad a = 1, 2, \dots, p \quad (4.4)$$

Let us now write out \tilde{A}^α explicitly (omitting the "one" over the "tweedle"). We shall start with:

$$T^{\alpha\beta}_{;\beta} = T^{\alpha\beta}_{;\beta} + \left\{ \begin{matrix} \alpha \\ \mu \nu \end{matrix} \right\} T^{\mu\nu}. \quad (4.5)$$

Introducing here (3.8) the first right hand expression becomes:

$$T^{\alpha\beta}_{;\beta} = T^{\alpha s}_{;s} + T^{\alpha 0}_{;0} = (m \overset{1}{\xi}{}^a \overset{1}{\xi}{}^s \delta)_{,s} + (m \overset{1}{\xi}{}^a \delta)_{,0} + \quad (4.6)$$

+ similar expressions concerning other particles.

We have:

$$\frac{dt}{ds} \int_{\Omega_{(3)}} (m \overset{1}{\xi}{}^a \overset{1}{\xi}{}^s \delta)_{,s} d_{(3)}x = \frac{dt}{ds} \int_{\Omega_{(3)}} T^{\alpha s}_{;s} d_{(3)}x = 0. \quad (4.7)$$

This we can see even without explicite calculations by changing the volume integral into a surface integral which must vanish because δ vanishes on the surface of $\Omega_{(3)}$. Therefore what remains of the integral of (4.6) is, because of (3.7),

$$\frac{dt}{ds} \int_{\Omega_{(3)}} T^{\alpha 0}_{;0} d_{(3)}x = \frac{dt}{ds} \frac{d}{dt} \left[\int_{\Omega_{(3)}} m \overset{1}{\xi}{}^a \delta d_{(3)}x \right] = \frac{dt}{ds} \frac{d}{dt} (m \overset{1}{\xi}{}^a) = (\mu \overset{1}{\xi}{}^a), \quad (4.8)$$

Thus because of (4.5), (3.6) and the last equation our equations of motion (4.2) become:

$$\mu \overset{1}{\xi}{}^a + \mu \overset{1}{\xi}{}^{a''} + \mu \left\{ \begin{matrix} \alpha \\ \mu \nu \end{matrix} \right\} \overset{1}{\xi}{}^{\mu} \overset{1}{\xi}{}^{\nu} = 0. \quad (4.9)$$

We shall show now that $\mu' = \frac{d\mu}{ds} = 0$, that is that the rest mass is constant. Indeed, let us assume, as we shall always do, that:

$$\left\{ \begin{matrix} \widetilde{\alpha} \\ \mu \nu \end{matrix} \right\} = \widetilde{g^{\alpha\sigma}} [\mu\nu, \varrho]. \quad (4.10)$$

Then, multiplying (4.9) by $\widetilde{g_{\alpha\sigma}} \xi^{\sigma}$, we have:

$$\mu' \widetilde{g_{\alpha\sigma}} \xi^{\alpha'} \xi^{\sigma'} + \mu \xi^{\alpha''} \xi^{\sigma'} \widetilde{g_{\alpha\sigma}} + \frac{1}{2} \mu \widetilde{g_{\mu\nu, \sigma}} \xi^{\mu'} \xi^{\nu'} \xi^{\sigma'} = 0. \quad (4.11)$$

Because of (2.9) this is equal to

$$\mu' + \frac{\mu}{2} \frac{d}{ds} (\widetilde{g_{\alpha\sigma}} \xi^{\alpha'} \xi^{\sigma'}) = 0; \quad \mu' = 0 \quad (4.12)$$

which changes one original equation (4.9) into

$$\boxed{\frac{1}{\mu} \widetilde{A^a} = \xi^{a''} + \left\{ \begin{matrix} \widetilde{\alpha} \\ \mu \nu \end{matrix} \right\} \xi^{\mu'} \xi^{\nu'} = 0} \quad (4.13)$$

Thus we see that the rest mass must be constant.

If not for the "tweedle" above the Christoffel symbol our equation (4.13) would be that of a geodetic line. For each particle we have four such equations of motion.

Let us put back t instead of s into (4.13) using (3.6). We then have:

$$\begin{aligned} \frac{dm}{dt} + m \left\{ \begin{matrix} \widetilde{o} \\ \alpha\beta \end{matrix} \right\} \dot{\xi}^{\alpha} \dot{\xi}^{\beta} &= 0. \\ \frac{d}{dt} (m \dot{\xi}^i) + m \left\{ \begin{matrix} \widetilde{s} \\ \alpha\beta \end{matrix} \right\} \dot{\xi}^{\alpha} \dot{\xi}^{\beta} &= 0; \quad (\dot{\xi}^o = 1). \end{aligned} \quad (4.14)$$

As we said before, we have $4p$ equations determining the ξ^k and the m^a 's. From the last equation in (4.14) and (3.8) it follows that:

$$\frac{m}{\mu} = \exp - \int \left\{ \begin{matrix} \widetilde{o} \\ \alpha\beta \end{matrix} \right\} \dot{\xi}^{\alpha} \dot{\xi}^{\beta} dt = \frac{dt}{ds}, \quad (4.15)$$

since

$$\frac{m}{m} = - \left\{ \begin{matrix} \widetilde{o} \\ \alpha\beta \end{matrix} \right\} \dot{\xi}^{\alpha} \dot{\xi}^{\beta} = \left(\log \frac{m}{\mu} \right)_{,0}. \quad (4.16)$$

From (4.15) we see that:

$$\widetilde{ds} = dt \exp \int \left\{ \begin{matrix} \widetilde{o} \\ \alpha\beta \end{matrix} \right\} \dot{\xi}^{\alpha} \dot{\xi}^{\beta} dt.$$

This connection between \widetilde{ds} and dt follows from the equations of motion. The normalisation is such that for $t = 0$, both dt and \widetilde{ds} are equal.

Substituting the values (4.16) for $\frac{m}{\mu}$ in the last three equations (4.14) we obtain the 3 p equations of motion for ξ^s :

$$\ddot{\xi}^s + \left\{ \begin{matrix} \widetilde{s} \\ \alpha \beta \end{matrix} \right\} \dot{\xi}^\alpha \dot{\xi}^\beta - \left\{ \begin{matrix} \widetilde{o} \\ \alpha \beta \end{matrix} \right\} \dot{\xi}^s \dot{\xi}^\alpha \dot{\xi}^\beta = 0; \quad \dot{\xi}^o = 1 \quad (4.17)$$

This equation, however, is not suitable for establishing a connection between the equations of motion and a variational principle. To do that let us rewrite the zero-th equation of (4.14) using (3.6)

$$\left(\log \frac{m}{\mu} \right)_{,0} = \left(\log \frac{dt}{\widetilde{ds}} \right)_{,0} = - \left\{ \begin{matrix} \widetilde{o} \\ \alpha \beta \end{matrix} \right\} \dot{\xi}^\alpha \dot{\xi}^\beta \quad (4.18)$$

or, since $\frac{dt}{\widetilde{ds}} = (\widetilde{g}_{\alpha\beta} \dot{\xi}^\alpha \dot{\xi}^\beta)^{-\frac{1}{2}}$:

$$\begin{aligned} \left[\log (\widetilde{g}_{\alpha\beta} \dot{\xi}^\alpha \dot{\xi}^\beta)^{-\frac{1}{2}} \right]_{,0} &= -\frac{1}{2} (\widetilde{g}_{\alpha\beta} \dot{\xi}^\alpha \dot{\xi}^\beta)^{-1} (\widetilde{g}_{\alpha\beta} \dot{\xi}^\alpha \dot{\xi}^\beta)_{,0} \\ &= - \left\{ \begin{matrix} \widetilde{o} \\ \alpha \beta \end{matrix} \right\} \dot{\xi}^\alpha \dot{\xi}^\beta. \end{aligned} \quad (4.19)$$

Therefore, the three equations (4.17) can also be written:

$$\ddot{\xi}^k + \left\{ \begin{matrix} \widetilde{k} \\ \alpha \beta \end{matrix} \right\} \dot{\xi}^\alpha \dot{\xi}^\beta - \frac{1}{2} (\widetilde{g}_{\alpha\beta} \dot{\xi}^\alpha \dot{\xi}^\beta)^{-1} (\widetilde{g}_{\alpha\beta} \dot{\xi}^\alpha \dot{\xi}^\beta)_{,0} \dot{\xi}^k = 0. \quad (4.20)$$

Let us write, for short:

$$\widetilde{\mathcal{L}} = (\widetilde{g}_{\alpha\beta} \dot{\xi}^\alpha \dot{\xi}^\beta)^{\frac{1}{2}}; \quad (\dot{\xi}^0 = 1). \quad (4.21)$$

Then we can rewrite (4.17) in the form

$$\ddot{\xi}^\sigma - (\log \widetilde{\mathcal{L}})_{,0} \dot{\xi}^\sigma + \left\{ \begin{matrix} \widetilde{\alpha} \\ \alpha \beta \end{matrix} \right\} \dot{\xi}^\alpha \dot{\xi}^\beta = 0. \quad (4.22)$$

Indeed the "zero" equation ($\sigma = 0$) gives the known equation (4.19). We shall multiply (4.22) by $\widetilde{g}_{\sigma\tau}$ and assume, from now on, that for $\widetilde{g}_{\alpha\beta}$ and their derivatives (2.17) is always valid, that is the „tweedled" product is equal to the products of the „tweedled" expressions. Then we have because of (2.9) the three equations of motion:

$$(\widetilde{g}_\sigma \dot{\xi}^\sigma)_{,0} - \widetilde{g}_{\sigma\tau} \dot{\xi}^\sigma (\log \widetilde{\mathcal{L}})_{,0} - \frac{1}{2} \widetilde{g}_{\alpha\beta,\tau} \dot{\xi}^\alpha \dot{\xi}^\beta = 0. \quad (4.23)$$

This form suggests the existence of a Lagrangian: $\widetilde{\mathcal{L}}$. We wish to see whether (4.23) is indeed equivalent to

$$\frac{d}{dt} \frac{\partial \widetilde{\mathcal{L}}}{\partial \dot{\xi}^k} - \frac{\partial \widetilde{\mathcal{L}}}{\partial \xi^k} = 0. \quad (4.24)$$

We find:

$$\begin{aligned} \frac{\partial \widetilde{\mathcal{L}}}{\partial \xi^k} &= \frac{1}{2} \widetilde{\mathcal{L}}^{-1} \widetilde{g}_{\alpha\beta, k} \dot{\xi}^\alpha \dot{\xi}^\beta \\ \frac{\partial \widetilde{\mathcal{L}}}{\partial \dot{\xi}^k} &= \frac{1}{2} \left(\frac{\partial \widetilde{g}_{\alpha\beta}}{\partial \dot{\xi}^k} \dot{\xi}^\alpha \dot{\xi}^\beta + 2 \widetilde{g}_{\alpha k} \dot{\xi}^\alpha \right) \widetilde{\mathcal{L}}^{-1} \\ \frac{d}{dt} \frac{\partial \widetilde{\mathcal{L}}}{\partial \dot{\xi}^k} &= \frac{1}{2} \left(\frac{\partial \widetilde{g}_{\alpha\beta}}{\partial \dot{\xi}^k} \dot{\xi}^\alpha \dot{\xi}^\beta \widetilde{\mathcal{L}}^{-1} \right)_{,0} + (\widetilde{g}_{\alpha k} \dot{\xi}^\alpha)_{,0} \widetilde{\mathcal{L}}^{-1} + \widetilde{g}_{\alpha k} \dot{\xi}^\alpha (\widetilde{\mathcal{L}}^{-1})_{,0}. \end{aligned} \quad (4.25)$$

Therefore the Lagrange equations (4.24) become:

$$\begin{aligned} (\widetilde{g}_{\alpha k} \dot{\xi}^\alpha)_{,0} - \widetilde{g}_{\alpha k} \dot{\xi}^\alpha (\log \widetilde{\mathcal{L}})_{,0} - \frac{1}{2} \widetilde{g}_{\alpha\beta, k} \dot{\xi}^\alpha \dot{\xi}^\beta + \\ + \frac{1}{2} \left(\frac{\partial \widetilde{g}_{\alpha\beta}}{\partial \dot{\xi}^k} \dot{\xi}^\alpha \dot{\xi}^\beta \widetilde{\mathcal{L}}^{-1} \right)_{,0} \widetilde{\mathcal{L}} = 0. \end{aligned} \quad (4.26)$$

Comparing (4.26) and (4.23) we see that $\widetilde{\mathcal{L}}$ is a Lagrangian if

$$\begin{aligned} \widetilde{g}_{\alpha\beta, k} &= \widetilde{g}_{\alpha\beta, k} = \frac{\partial \widetilde{g}_{\alpha\beta}}{\partial \dot{\xi}^k}, \\ \frac{\partial \widetilde{g}_{\alpha\beta}}{\partial \dot{\xi}^k} &= 0. \end{aligned} \quad (4.27)$$

This means: $\frac{d\widetilde{s}}{dt}$ is a Lagrangian if the relevant part of $g_{\alpha\beta}$ does not depend on $\dot{\xi}^s$ or $\dot{\xi}^s$.

We shall see later that the first condition, that is the independence of the relevant part of $g_{\alpha\beta}$ from $\dot{\xi}^s$ is not always fulfilled.

Thus (4.23) follows from a variational principle:

$$\delta \int_{-\infty}^{+\infty} \widetilde{\mathcal{L}} dt = \delta \int_{t=-\infty}^{+\infty} dt \left[\int_{\Omega(x)} d_{(3)}x \delta_{(3)}(x^s - \xi^s) g_{\alpha\beta} \dot{\xi}^\alpha \dot{\xi}^\beta \right]^{\frac{1}{2}} = 0, \quad (4.28)$$

if we treat $g_{\alpha\beta}$ as function of x^a only that is if, while varying the function under the integration sign, we ignore the possible dependence of $g_{\alpha\beta}$ on $\dot{\xi}$ and $\dot{\xi}$. Then, indeed, as can also be shown by a straightforward calculation (4.28) is equivalent to (4.23). If we call the result of such variation the "geodetic line" then (4.23) is the equation of a "geodetic line".

From what has been said it follows, that for a *test* particle — that is if $\mu \rightarrow 0$ and $g_{\alpha\beta}$ is not singular and does not depend on ξ or $\dot{\xi}$ — the equation (4.23) is that of a geodetic line and the “ \sim ” only means the substitution of $\dot{\xi}^i$ for x^i .

One can now show (Tulczyjew 1957) that the equations of motion follow from the field equations once we assume the linear dependence of $T^{\alpha\beta}$ from the δ 's, that is one can show that the form (3.6) for $t^{\alpha\beta}$ and (3.7) for $T^{\alpha\beta}$ is a consequence of the field equations.

Indeed from Bianchi's identities follows:

$$\frac{dt}{ds} \int_{\Omega(s)} \theta T^{\alpha\beta}_{;\beta} d_{(3)}x = 0, \quad (4.29)$$

where θ is an *arbitrary* function continuous on the worldline $\xi^i(t)$. Or, again omitting the “one” over the $t^{\alpha\beta}$'s and δ 's we can write:

$$\frac{dt}{ds} \int_{\Omega(s)} \theta (t^{\alpha\beta}_{;\beta})_{;\alpha} d_{(3)}x = 0. \quad (4.30)$$

Thus this equation leads to the following one:

$$\widetilde{A}^\alpha \theta + \widetilde{A}^{\alpha\beta} \widetilde{\theta}_{,\beta} = 0. \quad (4.31)$$

Let us start by calculating $\widetilde{A}^{\alpha\beta}$. Because

$$\delta_{,0} = -\delta_{,s} \dot{\xi}^s \quad (4.32)$$

we have:

$$\widetilde{\theta}_{,\beta} \widetilde{A}^{\alpha\beta} = \frac{dt}{ds} (-t^{\alpha s} + t^{\alpha\sigma} \dot{\xi}^s) \widetilde{\theta}_{,s}. \quad (4.33)$$

Since $\dot{\xi}^0 = 1$, this can be written:

$$\widetilde{\theta}_{,\beta} \widetilde{A}^{\alpha\beta} = \frac{dt}{ds} (-t^{\alpha\beta} + t^{\alpha\sigma} \dot{\xi}^\beta) \widetilde{\theta}_{,\beta}. \quad (4.34)$$

Because $\widetilde{\theta}_{,\beta}$ is arbitrary, we have:

$$\widetilde{A}^{\alpha\beta} = \frac{dt}{ds} (-t^{\alpha\beta} + t^{\alpha\sigma} \dot{\xi}^\beta). \quad (4.35)$$

Putting here $\alpha = 0$ we find:

$$t^{0\beta} = t^{00} \dot{\xi}^\beta, \quad (4.36)$$

therefore generally

$$t^{a\beta} = t^{00} \dot{\xi}^a \dot{\xi}^\beta = m \dot{\xi}^a \dot{\xi}^\beta \quad (4.37)$$

which is the proof of our theorem. Obviously $\tilde{A}^a = 0$ gives the equations of motion.

§ 5. The approximation method

We shall solve the field equations and formulate the equations of motion explicitly by means of the approximation method which we shall presently describe.

Let us assume a function developed into a power series in the parameter $\lambda = \frac{1}{c}$ (for the moment we may assume c arbitrary and not equal to "one").

$$\varphi = \varphi_0 + \varphi_1 + \varphi_2 + \dots \quad (5.1)$$

The indices written below indicate the order of λ absorbed by the φ 's.

If the function $\varphi(x^\mu)$ varies rapidly in space but slowly with x^0 then we are justified in not treating all its derivatives in the same manner. The derivatives with respect to x^0 will be of a higher order than the space derivatives. We can formalise this procedure by assuming:

$$\frac{\partial}{\partial x^0} (\varphi) = \varphi_{i,0} = \varphi_{i+1,0} \quad (5.2)$$

that is: differentiation with respect to x^0 raises the power by λ , absorbed by the φ 's, by one.

The problem is now with what order we should start the power development of the quantities appearing in the field equations?

The quantity ξ^s will start (by an obvious convention) with the order "zero". It will be an unknown quantity determined by the equations of motion; we shall for the moment not develop into a power series. Thus $\dot{\xi}^s$ will be of the order "one" and $\ddot{\xi}^s$ of the order "two".

From this follows:

$$m = m_2 + m_4 + m_6 + \dots \quad (5.3)$$

To begin with m is not pure convention. Indeed in the Newtonian approximation, which we hope to obtain, we have in the chosen units

$$\text{mass} \times \text{acceleration} = \frac{\text{mass} \times \text{mass}}{(\text{distance})^2}$$

Since the acceleration is of order two, the orders of both sides will be equal, only if the order of mass is also "two".

In all the power developments we shall take into account only even (as in $\overset{a}{m}$), or only odd powers of λ . (Infeld 1938).

Thus because of the order with which we start $\overset{a}{m}$ and λ we have:

$$\begin{aligned} T^{00} &= T^{00}_2 + T^{00}_4 + T^{00}_6 + \dots \\ T^{0m} &= T^{0m}_3 + T^{0m}_5 + T^{0m}_7 + \dots \\ T^{mn} &= T^{mn}_6 + T^{mn}_8 + \dots \end{aligned} \quad (5.4)$$

Now as to the g 's we write:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}; \quad g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu} \quad (5.5)$$

From the gravitational equation follows:

$$R_{\alpha\beta} = -8\pi \left(T_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} T \right) \quad (5.6)$$

where

$$\begin{aligned} R_{\alpha\beta} &= R_{\alpha\beta} \sqrt{-g} \\ R_{\mu\nu} &= - \left\{ \begin{matrix} \varrho \\ \mu \nu \end{matrix} \right\}_{,\varrho} + \left\{ \begin{matrix} \varrho \\ \mu \varrho \end{matrix} \right\}_{,\nu} + \left\{ \begin{matrix} \varrho \\ \mu \sigma \end{matrix} \right\} \left\{ \begin{matrix} \sigma \\ \varrho \nu \end{matrix} \right\} - \left\{ \begin{matrix} \varrho \\ \mu \nu \end{matrix} \right\} \left\{ \begin{matrix} \sigma \\ \varrho \sigma \end{matrix} \right\}. \end{aligned} \quad (5.7)$$

From the right hand side of equation (5.6) follows: R_{00} and R_{mn} start with the order "2" and R_{0m} with the order "3". The lowest order expressions on the left hand side are:

$$\begin{aligned} \text{for } R_{00}: & -\frac{1}{2} h_{00,ss} \\ \text{for } R_{0m}: & -\frac{1}{2} h_{0m,ss} + \frac{1}{2} h_{0s,ms} + \frac{1}{2} h_{ms,0s} - \frac{1}{2} h_{ss,m0} \\ \text{for } R_{mn}: & -\frac{1}{2} h_{mn,ss} + \frac{1}{2} h_{ms,ns} + \frac{1}{2} h_{ns,ms} + \frac{1}{2} h_{00,mn} - \frac{1}{2} h_{ss,mn}. \end{aligned} \quad (5.8)$$

Therefore

$$\begin{aligned} h_{00} &= h_{00}_2 + h_{00}_4 + \dots \\ h_{0m} &= h_{0m}_3 + h_{0m}_5 + \dots \\ h_{mn} &= h_{mn}_2 + h_{mn}_4 + \dots \end{aligned} \quad (5.9)$$

All the functions that will appear later are obtained from the h 's by summation, multiplication, differentiation. Thus to every component, the following rule applies throughout: any component having an odd (even) number of zero suffixes will have only odd (even) powers of λ in its expansion.

§ 6. The Newtonian equations of motion

We shall try to find the equations of motion in the lowest (Newtonian) approximation. We shall do it in such a way as to make the generalisation to the post-Newtonian approximation as simple as possible.

Because of (5.6) and (5.8), the field equations of the lowest order are in h_{00} :

$$-\frac{1}{2} h_{00,ss} = -8\pi \left(T_{2}^{00} - \frac{1}{2} T_{2}^{00} \right) = -4\pi T_{2}^{00} = -4\pi \sum_{a=1}^p m^a \delta^a, \quad (6.1)$$

or:

$$h_{00,ss} = 8\pi \sum_{a=1}^p m^a \delta^a. \quad (6.2)$$

As the solution of this equation we take the Newtonian field, that is:

$$h_{00} = -2 \frac{1}{2} m r^{-1} - 2 \frac{2}{2} m r^{-1} - \dots \quad (6.3)$$

where

$$r^a = (x^s - \xi^s) (x^s - \xi^s). \quad (6.4)$$

We shall also write briefly:

$$h_{00} = \varphi \quad (6.5)$$

and in the two body case, which we shall assume for the sake of simplicity³

$$\left. \begin{aligned} \varphi &= f + g \\ f &= -2 \frac{1}{2} m r^{-1}; & g &= -2 \frac{2}{2} m r^{-1} \\ \xi &= \xi; & \xi &= \eta \end{aligned} \right\} \quad (6.6)$$

It is also worth remarking that because of (5.5) and

$$g^{a\alpha} g_{\beta\alpha} = \delta^a_{\beta}, \quad (6.7)$$

we have

$$h_{00} = -\varphi \quad (6.8)$$

³ There is, I hope, no danger in confusing this g in (6.6) with $g = |g_{\alpha\beta}|$.

We see that the relevant part of φ , that is g does not depend on ξ ; therefore we have:

$$\widetilde{h_{00s}} = \widetilde{h_{00s}} = \widetilde{g_{ss}} = \frac{\partial \widetilde{g}}{\partial \xi^s} = g_{, \xi^s} \quad (6.9)$$

There was no reason to assume that m is a constant. This follows, however, from (4.15) or the first equation in (4.14).

Because of (6.9) and because in the lowest order we do not have any products of the h 's, the conditions (4.10) and (4.27) are satisfied, this means that the path is a "geodetic line" and can be deduced from a Lagrangian:

$$\widetilde{\mathcal{L}} = (\widetilde{g_{\alpha\beta}} \dot{\xi}^\alpha \dot{\xi}^\beta)^{\frac{1}{2}}; \quad \widetilde{\mathcal{L}}_2 = (1 - \dot{\xi}^s \dot{\xi}^s + \widetilde{g})^{\frac{1}{2}} \quad (6.10)$$

$$\boxed{\begin{aligned} \mathcal{L}_2 &= -\frac{1}{2} \dot{\xi}^s \dot{\xi}^s + \frac{1}{2} \widetilde{g} = -\frac{1}{2} \dot{\xi}^s \dot{\xi}^s - \frac{m}{r} \\ r^2 &= (\xi^s - \eta^s)(\xi^s - \eta^s) \end{aligned}} \quad (6.11)$$

Thus the equations of motion, up to the second order, are:

$$\frac{d}{dt} \frac{\partial \widetilde{\mathcal{L}}}{\partial \dot{\xi}^s} - \frac{\partial \widetilde{\mathcal{L}}}{\partial \xi^s} = 0. \quad (6.12)$$

that is:

$$\ddot{\xi}^s = \frac{\partial}{\partial \xi^s} \left(\frac{m}{r} \right). \quad (6.13)$$

The Lagrangian $\widetilde{\mathcal{L}}^{**}$ for both particles is

$$\boxed{\widetilde{\mathcal{L}}^{**} = -\frac{1}{2} m^1 \dot{\xi}^s \dot{\xi}^s - \frac{1}{2} m^2 \dot{\eta}^s \dot{\eta}^s - \frac{m^1 m^2}{r}} \quad (6.14)$$

The Newtonian Lagrangian for both particles is of the fourth order! One more remark

Since $\frac{m}{\mu} = \frac{dt}{ds}$, we have here:

$$\boxed{\begin{aligned} m &= \mu \\ \frac{1}{m} &= \frac{1}{2} m^1 \dot{\xi}^s \dot{\xi}^s + \frac{m^1 m^2}{r} \end{aligned}} \quad (6.15)$$

We can also find the equations of motion directly by going back to (4.17). We then have simply:

$$\ddot{\xi}^s + \left\{ \begin{matrix} s \\ o o \end{matrix} \right\} = 0 \quad (6.16)$$

or

$$\ddot{\xi}^s - [00, s] = \ddot{\xi}^s + \frac{1}{2} \widetilde{g}_{ss} = 0. \quad (6.17)$$

which is identical with (6.13).

§ 7. Transition to the next approximation

To find the equations of motion up to the fourth order besides h_{00} we must also know

$$h_{mn}^{(2)}; h_{om}^{(3)}; h_{00}^{(4)}. \quad (7.1)$$

Thus we must proceed to calculate these expressions. As far as the first two are concerned, the task is easy. The left hand side of the corresponding equations is written out in (5.8) and the right hand side is given by (5.6) and it is:

$$\begin{aligned} \text{for } mn: & -4\pi \delta_{mn}^{(2)} \left(m \delta^{(1)} + m \delta^{(2)} \right) \\ \text{for } om: & 8\pi \left(m \xi^m \delta^{(1)} + m \eta^m \delta^{(2)} \right) \end{aligned} \quad (7.2)$$

Therefore, for $h_{mn}^{(2)}$ we have the equations:

$$-\frac{1}{2} h_{mn,ss}^{(2)} + \frac{1}{2} h_{ms,ns}^{(2)} + \frac{1}{2} h_{ns,ms}^{(2)} - \frac{1}{2} h_{ss,mn}^{(2)} + \frac{1}{2} \varphi_{,mn} = -\frac{1}{2} \delta_{mn} \varphi_{,ss}. \quad (7.3)$$

We are looking for a solution of a Newtonian character. Such a solution is:

$$h_{mn}^{(2)} = \delta_{mn} \varphi. \quad (7.4)$$

The choice of these Newtonian solutions for $h_{00}^{(2)}$ and $h_{mn}^{(2)}$ must be regarded as part of our approximation procedure. More about this later in § 9.

The next step is to calculate $h_{om}^{(3)}$. Again (5.8) and (7.2) gives:

$$-\frac{1}{2} h_{on,ss}^{(3)} + \frac{1}{2} h_{os,ns}^{(3)} + \frac{1}{2} h_{ns,os}^{(3)} - \frac{1}{2} h_{ss,no}^{(3)} = 8\pi \left(m \xi^n \delta^{(1)} + m \eta^n \delta^{(2)} \right) \quad (7.5)$$

Introducing (7.4) into the last equation we have:

$$-\frac{1}{2} h_{on,ss}^{(3)} + \frac{1}{2} h_{os,ns}^{(3)} - \varphi_{,no} = 8\pi \left(m \xi^n \delta^{(1)} + m \eta^n \delta^{(2)} \right). \quad (7.6)$$

The solution which we shall adopt here is

$$h_{on} = -2f\dot{\xi}^n - 2g\dot{\eta}^n. \quad (7.7)$$

This is not the only possible solution. About a more general solution and its influence upon the equations of motion we shall say something in §9.

The calculation of h_{00} is more troublesome. We shall give it in Appendix B. Here we shall quote the relevant expressions of h_{00} , that is those that give a contribution to \widetilde{h}_{00} and $\widetilde{h}_{00,\alpha}$. They are:

$$h_{00} \sim 2 \binom{2}{m}^2 \binom{2}{r}^{-2} - 3 \eta^s \eta^s m \binom{2}{r}^{-1} - m r_{,00} + 2 m m \binom{2}{r r}^{-1}. \quad (7.8)$$

Now we should like to see whether these expressions for h_{mn} , h_{0n} , h_{00} are such that they make the equations of motion equivalent to those of a „geodetic line”, that is, if the conditions

$$\left\{ \begin{matrix} \widetilde{\alpha} \\ \beta \varrho \end{matrix} \right\} = \widetilde{g}^{\alpha\sigma} [\beta \varrho, \sigma] \quad (7.9)$$

$$\widetilde{g}_{\alpha\beta,s} = \widetilde{g}_{\alpha\beta,s}; \quad \frac{\partial \widetilde{g}_{\alpha\beta}}{\partial \xi^r} = 0 \quad (7.10)$$

are satisfied. Both conditions are satisfied for h_{mn} and h_{0n} . Indeed this expressions have a singularity of the order $\frac{1}{r}$, therefore an odd singularity, therefore for them

(7.9) is satisfied. Furthermore in their relevant part neither ξ nor $\dot{\xi}$ appears explicitly; therefore for them (7.10) is satisfied. But this is not true of h_{00} . It has an even

singularity because of the appearance of $\binom{1}{r}^{-2}$. Yet this does not matter since h_{00} appears only *linearly* in the equations of motion. Therefore we can disregard the condition (7.9) for h_{00} . But is the condition (7.10) satisfied for h_{00} ? It is certainly satisfied for the first two expressions, that is for $2 \binom{2}{m}^2 \binom{2}{r}^{-2} - 3 \eta^s \eta^s m \binom{2}{r}^{-1}$, since neither of these two expressions depends on ξ , $\dot{\xi}$. We must be more careful with the third expression:

$$-m r_{,00} = -m r_{, \eta^s \eta^s} r - m r_{, \eta^s \eta^s} r. \quad (7.11)$$

The first expression on the right hand side does not depend on ξ , $\dot{\xi}$, but the second expression contains η^s . Here we may introduce for $m \eta^s = -m \dot{\xi}^s$, the Newtonian

value since the mistake will be of order "6". Thus we can put:

$$\alpha = - \overset{2}{m} \overset{2}{r, \eta^s} \overset{..}{\eta^s} = \overset{1}{m} \overset{2}{m} \overset{2}{r, \eta^s} \left(\frac{1}{r} \right)_{, \xi^s} = \left[\overset{1}{m} \overset{2}{m} (x^s - \eta^s) \left(\frac{2}{r} \right)^{-1} \right] (x^s - \eta^s) r^{-3} \quad (7.12)$$

This expression, differentiated with respect to x^m gives zero at the point $x^s = \xi^s$. Thus

$$\widetilde{\alpha}_{,m} = 0 \quad (7.13)$$

But

$$\widetilde{\alpha}_{,m} = \frac{\partial \widetilde{\alpha}}{\partial \xi^m} = 2 \overset{1}{m} \overset{2}{m} \left(\frac{1}{r} \right)_{, \xi^m}.$$

A similar situation occurs in the last expression in h_{00} :

$$\beta = 2 \overset{1}{m} \overset{2}{m} \left(\frac{2}{r} \right)^{-1} \quad (7.14)$$

We have:

$$\begin{aligned} \widetilde{\beta}_{,s} &= 2 \overset{1}{m} \overset{2}{m} r^{-1} \left[\left(\frac{2}{r} \right)^{-1} \right]_{,s} = 2 \overset{1}{m} \overset{2}{m} r^{-1} (r^{-1})_{, \xi^m} \\ \widetilde{\beta}_{,s} &= 2 \overset{1}{m} \overset{2}{m} (r^{-2})_{, \xi^m} \end{aligned} \quad (7.15)$$

Therefore:

$$\widetilde{\beta}_{,s} = 2 \widetilde{\beta}_{,s} \quad (7.16)$$

We can, however, easily find an auxiliary field which we shall denote by h_{00}^* such, that:

$$\widetilde{h_{00,m}^*} = \widetilde{h_{00,m}} \quad (7.17)$$

From this it follows that such $\widetilde{h_{00}^*}$ is:

$$\widetilde{h_{00}^*} = 2 \overset{2}{(m)^2} r^{-2} - 3 \overset{.}{\eta^s} \overset{.}{\eta^s} \overset{2}{m} r^{-1} - \overset{2}{m} r_{, \xi^s \xi^r} \overset{.}{\eta^s} \overset{.}{\eta^r} + \overset{1}{m} \overset{2}{m} r^{-2} \quad (7.18)$$

§ 8. The post Newtonian equations of motion

If we look at the general equations of motion (4.17) we see that — up to the fourth order, h_{00} will appear only once, that is in the expression $\frac{1}{2} h_{00,m}$. This means, because of (7.17) that a Lagrangian up to the fourth order exists⁴:

$$\widetilde{\mathcal{L}}^* = \frac{d\widetilde{s}^*}{dt} \quad (8.1)$$

⁴ Compare Fichtenholz (1954) where the Lagrangian is found mechanically from the explicit equations of motion and not as here, the equation of motion from the Lagrangian.

with

$$\left(\frac{d\tilde{s}^*}{dt}\right)^2 = \tilde{g}_{\alpha\beta}^* \dot{\xi}^\alpha \dot{\xi}^\beta \quad (8.2)$$

in which only $\tilde{g}_{00}^* \neq \tilde{g}_{00}$ and

$$\tilde{g}_{00}^* = \eta_{00} + h_{00} + h_{00}^* = 1 + \varphi + h_{00}^* \quad (8.3)$$

Therefore:

$$\tilde{\mathcal{L}}^* = (1 + \tilde{\varphi} + \tilde{h}_{00}^* - \dot{\xi}^s \dot{\xi}^s + \tilde{\varphi} \dot{\xi}^s \dot{\xi}^s + h_{s0} \dot{\xi}^s)^{\frac{1}{2}}. \quad (8.4)$$

Here we introduce the values:

$$\begin{aligned} \tilde{\varphi} &= -\frac{2}{2} m r^{-1}; & \tilde{h}_{0s}^* &= 4 \frac{2}{3} m r^{-1} \dot{\eta}^s \\ \tilde{h}_{00}^* &= 2 \frac{2}{4} (m)^2 r^{-2} - 3 \dot{\eta}^s \dot{\eta}^s m r^{-1} - m r_{,\xi^s \xi^r} \dot{\eta}^s \dot{\eta}^r + m m r^{-1}. \end{aligned} \quad (8.5)$$

Remembering that up to the fourth order

$$(1 + a + \frac{a^2}{4})^{\frac{1}{2}} = 1 + \frac{1}{2} (a + \frac{a^2}{4}) - \frac{1}{8} a^2 \quad (8.6)$$

we obtain from (8.4) and (8.5) the Lagrangian:

$$\begin{aligned} \tilde{\mathcal{L}}^* &= \tilde{\mathcal{L}}^* + \tilde{\mathcal{L}}^* = -\frac{1}{2} \dot{\xi}^s \dot{\xi}^s - m r^{-1} - \frac{3}{2} m^2 r^{-1} (\dot{\xi}^s \dot{\xi}^s + \dot{\eta}^s \dot{\eta}^s) + \\ &+ 4 \frac{2}{m r^{-1}} \dot{\xi}^s \dot{\eta}^s - \frac{1}{8} (\dot{\xi}^s \dot{\xi}^s)^2 + \frac{1}{2} \frac{m^2 (m + m)}{r^2} - \frac{1}{2} m r_{,\xi^s \xi^r} \dot{\eta}^s \dot{\eta}^r. \end{aligned} \quad (8.7)$$

This is supposed to be the Lagrangian for the first particle. But we wish to find the Lagrangian for *both* particles. This means: *first*: a Lagrangian which gives the same equations of motion as the Lagrangian (8.7); *secondly*: a Lagrangian which is invariant with respect to a transformation, changing:

$$\frac{1}{m\xi} \rightarrow \frac{2}{m\eta}. \quad (8.8)$$

Let us multiply $\tilde{\mathcal{L}}^*$ in (8.7) by m . Then the equations of motion will, of course, be the same. Let us add:

$$-\frac{1}{2} m^2 \dot{\eta}^s \dot{\eta}^s - \frac{1}{8} m^2 (\dot{\eta}^s \dot{\eta}^s)^2. \quad (8.9)$$

Then, with this addition, obviously, the equations of motion for the first particle will still be the same since there is no contribution from (8.9). Now with these changes,

the only expression not invariant with the respect to the change (8.8) will be the last one in (8.7), that is:

$$-\frac{1}{2} \frac{1}{m} \frac{2}{m} r_{,\xi^s \xi^r} \dot{\xi}^s \dot{\eta}^r \quad (8.10)$$

But instead of it we can write:

$$\frac{1}{2} \frac{1}{m} \frac{2}{m} r_{,\xi^s \eta^r} \dot{\xi}^s \dot{\eta}^r \quad (8.11)$$

which is invariant with respect to the change (8.8), and gives the same contributions to the equations of motion.

Indeed from (8.10) we have the following addition to the equations of motion of the first particle:

$$\frac{1}{2} \frac{1}{m} \frac{2}{m} r_{,\xi^s \xi^r \xi^n} \dot{\eta}^s \dot{\eta}^r = -\frac{1}{2} \frac{1}{m} \frac{2}{m} r_{,\xi^s \eta^r \xi^n} \dot{\eta}^s \dot{\eta}^r. \quad (8.12)$$

From (8.11) we have the addition:

$$\begin{aligned} & \frac{1}{2} \frac{1}{m} \frac{2}{m} (r_{,\xi^n \eta^r} \dot{\eta}^r)_{,0} - \frac{1}{2} \frac{1}{m} \frac{2}{m} r_{,\xi^s \eta^r \xi^n} \dot{\xi}^s \dot{\eta}^r = \\ & \frac{1}{2} \frac{1}{m} \frac{2}{m} (r_{,\xi^n \eta^r \xi^s} \dot{\eta}^r \dot{\xi}^s + r_{,\xi^n \eta^r \eta^s} \dot{\eta}^r \dot{\eta}^s + \\ & + r_{,\xi^n \eta^r} \ddot{\eta}^r - r_{,\xi^s \eta^r \xi^n} \dot{\xi}^s \dot{\eta}^r) = \frac{1}{2} \frac{1}{m} \frac{2}{m} (r_{,\xi^n \eta^r \eta^s} \dot{\eta}^r \dot{\eta}^s + r_{,\xi^n \eta^r} \ddot{\eta}^r). \end{aligned} \quad (8.13)$$

But the last expression in (8.13) equals:

$$\frac{1}{2} \frac{1}{m} \frac{2}{m} r_{,\xi^n \eta^r} \ddot{\eta}^r = \frac{1}{2} \frac{1}{m} \frac{2}{m} \left(-\frac{\delta_{nr}}{r} + \frac{(\xi^n - \eta^n)(\xi^r - \eta^r)}{r^3} \right) \ddot{\eta}^r \quad (8.14)$$

Since

$$\ddot{\eta}^r = \frac{1}{r^3} m (\xi^r - \eta^r) \quad (8.15)$$

we see that (8.14) vanishes and, therefore, (8.10) and (8.11) give the same contributions to the equations of motion.

Therefore, if we call the final Lagrangian for two particles $\tilde{\mathcal{L}}^{**}$, we have:

$$\begin{aligned} \tilde{\mathcal{L}}^{**} &= \tilde{\mathcal{L}}^{**}_4 + \tilde{\mathcal{L}}^{**}_6 = -\frac{1}{2} \frac{1}{m} \dot{\xi}^s \dot{\xi}^s - \frac{1}{2} \frac{2}{m} \dot{\eta}^s \dot{\eta}^s - \frac{1}{r} \frac{1}{m} \frac{2}{m} \\ & - \frac{3}{2} \frac{1}{m} \frac{2}{m} \frac{1}{r} (\dot{\xi}^s \dot{\xi}^s + \dot{\eta}^s \dot{\eta}^s) + 4 \frac{1}{m} \frac{2}{m} \frac{1}{r} \dot{\xi}^s \dot{\eta}^s - \frac{1}{8} \frac{1}{m} (\dot{\xi}^s \dot{\xi}^s)^2 - \\ & - \frac{1}{8} \frac{1}{m} (\dot{\eta}^s \dot{\eta}^s)^2 + \frac{1}{2} \frac{1}{m} \frac{2}{m} \frac{1}{r^2} (m + m) + \frac{1}{2} \frac{1}{m} \frac{2}{m} r_{,\xi^s \eta^r} \dot{\xi}^s \dot{\eta}^r \end{aligned} \quad (8.16)$$

By a straightforward calculation we find from this Lagrangian the equations of motion of the first particle:

$$\ddot{\xi}^n - m \left(\frac{1}{r} \right)_{,\xi^n} = m \left\{ \left[\dot{\xi}^s \dot{\xi}^s + \frac{3}{2} \dot{\eta}^s \dot{\eta}^s - 4 \dot{\xi}^s \dot{\eta}^s - 4 \frac{m}{r} - 5 \frac{m}{r} \right] \left(\frac{1}{r} \right)_{,\xi^n} + \right. \\ \left. + [4 \dot{\xi}^s (\dot{\eta}^n - \dot{\xi}^n) + 3 \dot{\xi}^n \dot{\eta}^s - 4 \dot{\eta}^n \dot{\eta}^s] \left(\frac{1}{r} \right)_{,\xi^s} + \frac{1}{2} r_{,\xi^s \xi^s \eta^s} \dot{\eta}^s \dot{\eta}^s \right\}. \quad (0.17)$$

The equations of motion for the other particle are obtained by replacing

$$\begin{matrix} 1 & 2 \\ m, & m, \end{matrix} \xi, \eta \quad \text{by} \quad \begin{matrix} 2 & 1 \\ m, & m, \end{matrix} \eta, \xi \quad (8.18)$$

respectively.

The generalisation of this result to p particles is almost trivial, if we take into account the changes caused by the addition of these particles in \widetilde{h}_{00}^* and outlined at the end of the Appendix B. These additional expressions are due, say, in the case of three particles, to the interaction between the second and third particles; that is in the equations of motion for the first particle they will give a contribution proportional

$\begin{matrix} 1 & 2 & 3 \\ m & m & m. \end{matrix}$ These expressions appear in $\widetilde{\mathcal{L}}^{**}$ from two sources: from \widetilde{h}_{00}^* and from $\widetilde{\varphi}^2$ in (8.6). If we now denote the "distance" from the a -th to the b -th particle by:

$$r^{(ab)} = [(\dot{\xi}^s - \dot{\xi}^s)^a (\dot{\xi}^s - \dot{\xi}^s)^b]^{1/2}, \quad (8.19)$$

then we have the Lagrangian for the p particles:

$$\begin{aligned} \widetilde{\mathcal{L}}^{**} &= \widetilde{\mathcal{L}}_4^{**} + \widetilde{\mathcal{L}}_6^{**} = -\frac{1}{2} \sum_{a=1}^p m^a \dot{\xi}^s \dot{\xi}^s - \frac{1}{2} \sum_{\substack{a,b=1 \\ a \neq b}}^p m^a m^b \left(\frac{1}{r} \right)^{(ab)} - \\ &- \frac{3}{4} \sum_{\substack{a,b=1 \\ a \neq b}}^p m^a m^b \left(\frac{1}{r} \right)^{(ab)} (\dot{\xi}^s \dot{\xi}^s + \dot{\xi}^s \dot{\xi}^s) + 2 \sum_{\substack{a,b=1 \\ a \neq b}}^p m^a m^b \left(\frac{1}{r} \right)^{(ab)} \dot{\xi}^s \dot{\xi}^s - \\ &- \frac{1}{8} \sum_{a=1}^p m^a (\dot{\xi}^s \dot{\xi}^s)^2 + \frac{1}{4} \sum_{\substack{a,b=1 \\ a \neq b}}^p m^a m^b (m^a + m^b) \left(\frac{1}{r} \right)^{(ab)} + \\ &+ \sum_{\substack{a,b=1 \\ a \neq b}}^p m^a m^b r_{,\xi^s \xi^s} \dot{\xi}^s \dot{\xi}^s + \frac{1}{6} \sum_{\substack{a,b,c=1 \\ a \neq b \neq c}}^p m^a m^b m^c \left[\left(\frac{1}{r} \right)^{(ab)(ac)} + \left(\frac{1}{r} \right)^{(bc)(ba)} + \left(\frac{1}{r} \right)^{(ca)(cb)} \right]. \end{aligned} \quad (8.20)$$

In the case of two particles the Lagrangian (8.20) reduces to (8.16). The only new expression appearing in (8.20) is the last one; in the case of three particles it is equal to:

$$\frac{1}{m} \left(\frac{1}{2} \widetilde{S}_{00} - \frac{1}{2} \widetilde{g} \widetilde{k} \right)$$

if we accept the notation of Appendix B where \widetilde{S}_{00} is the change in h_{00} caused by the interaction of the second and third particle and k is for the third particle what f and g are for the first and second. Thus $-\frac{1}{4} \widetilde{g} \widetilde{k}$ is the contribution to the Lagrangian of the interaction between the second and third particles coming from $-\frac{1}{8} \widetilde{\varphi}^2$ in (8.6).

§ 9. On the choice of the coordinate-system

The harmonic coordinate condition is

$$\sqrt{-g} g^{\mu\nu}{}_{, \nu} = 0, \quad (9.1)$$

which in our case means

$$h_{2, n}^{mn} = 0; \quad h_{3, m}^{0m} + h_{2, 0}^{00} = 0. \quad (9.2)$$

None of this conditions are fulfilled in our coordinate-system.

The values for h_{00} and h_{mn} accepted by us here were the Newtonian values. Our convention is that their choice characterised our approximation procedure. Yet, with some justification, this approach may be regarded as too formal. Indeed, instead of our values for h_{mn} we could have chosen

$$h'_{2, mn} = h_{2, mn} + a_{2, n} + a_{n, m}$$

the a 's being arbitrary functions. This change from $h_{2, mn}$ to $h'_{2, mn}$ could also be induced by a change in a coordinate-system that does not disturb our approximation procedure. What is the physical meaning of the choice $a_m = 0$? It means that we assume the existence of a coordinate-system in which each of the two bodies reveals its spherical symmetry; a coordinate-system in which for $m \rightarrow 0$, $\xi \rightarrow 0$, the field goes over into that defined by the Schwarzschild solution in an isotropic coordinate-system. The choice of such a coordinate-system is implicitly assumed by our approximation procedure. However this choice of $a_m = 0$ refers only to the beginning of our approximation procedure; therefore it does not refer to h_{0m} . Indeed if we replace h_{0m} by

$$h'_{3, 0m} = h_{3, 0m} + a_{0, m} \quad (9.3)$$

a_0 being an arbitrary function of x^a , then the equation (7.5) is fulfilled just as well.

This change in h_{0m} can be induced by a change in a coordinate-system which does not disturb our approximation procedure. Such a change also induces a simple change in h_{00} (see Appendix B)

$$h_{00}^* = h_{00} + 2 a_{0,0} \quad (9.4)$$

Therefore it would seem that the Lagrangian and with it the equations of motion would change. The expressions that change in the Lagrangian (8.4) are:

$$\widetilde{h}_{0s} \dot{\xi}^s + \frac{1}{2} \widetilde{h}_{00}^* \quad (9.5)$$

and the change induced by them in the Lagrangian is, because of (9.3) and (9.4):

$$\Delta \widetilde{\mathcal{L}}^* = \widetilde{a}_{0,s} \dot{\xi}^s + \widetilde{a}_{0,0} = \frac{d\widetilde{a}_0}{dt}. \quad (9.6)$$

Therefore

$$\delta \int_{t_1}^{t_2} \Delta \widetilde{\mathcal{L}}^* dt = \delta(a_0) \Big|_{t_1}^{t_2} = 0. \quad (9.7)$$

This means: the equations of motion are uniquely determined up to the fourth approximation by the field equations and by our approximation procedure. Neither the harmonic coordinate condition nor any other coordinate condition played any role in our derivation of the equation of motion.

§ 10. The general theory

Let us now formulate the general theory⁵, according to which we proceeded in our special case, in which we found the equations of motion of the second and fourth order. Of course, such a general theory is of little practical value since there would be hardly any physical meaning in developing the calculations one step further. Moreover, it seems — and we shall discuss this later — that by the proper choice of the coordinate-system we can annihilate all contributions to the equations of motion beyond the fourth order. Yet from the formal point of view it is important to know that the procedure can be pushed as far as we wish. Of course we do not know anything about its convergence.

Before we formulate the general theory let us recall what has been done here. We had the Newtonian equations of motion:

$$\frac{dt}{ds} \int_3 T^{0\nu}_{;\nu} d_{(3)} x = \widetilde{A}^0_3 = 0; \quad \frac{dt}{ds} \int_4 T^{n\nu}_{;\nu} d_{(3)} x = \widetilde{A}^n_4 = 0. \quad (10.1)$$

⁵ The ideas presented here are a few years old. Scheidegger (1953) refers to them in § 5 stating that they were suggested by me. I found a more explicit formulation of similar ideas in a thesis by Rameswararao (1955).

Since m appears as a factor we called these equations (after dividing them by m) the equations of the *second* order. But here, since it is multiplied by m it appears as an equation of the fourth order. For this section therefore, let us rename the order of the equations of motion calling the *Newtonian* equations of motion those of the *fourth* order and the post-Newtonian equations those of the *sixth* order. Let us also put generally:

$$\frac{dt}{ds} \int T^{ov}_{2n-1} d_{(3)} x = \widetilde{A}^0_{2n-1}; \quad \frac{dt}{ds} \int T^{nv}_{2n} d_{(3)} x = \widetilde{A}^n_{2n}. \quad (10.2)$$

Thus for our post-Newtonian equations of motion we have:

$$\widetilde{A}^0_3 + \widetilde{A}^0_5 = 0; \quad \widetilde{A}^n_2 + \widetilde{A}^n_4 = 0 \quad (10.3)$$

These equations gave us:

$$m = m^a_2 + m^a_4; \quad \xi^s = \xi^s_0 + \xi^s_2, \quad (10.4)$$

where ξ^s_0 is the motion in the Newtonian approximation. To find these equations explicitly we used the Newtonian equation of motion in \widetilde{A}^m_6 since the use of ξ^s_2 instead of ξ^s_0 would give a contribution of the 8-th order to the equations of motion. We express this idea in symbols and write instead of (10.3):

$$\begin{aligned} \widetilde{A}^0_3 (\xi^k_0 + \xi^k_2) + \widetilde{A}^0_5 (\xi^k_0) &= 0, \\ \widetilde{A}^m_4 (\xi^k_0 + \xi^k_2) + \widetilde{A}^m_6 (\xi^k_0) &= 0. \end{aligned} \quad (10.5)$$

Thus $\widetilde{A}^m_4 (\xi^k_0 + \xi^k_2)$ also gives a contribution of the order „six“. The field was solved so as to obtain (10.4) in:

$$\begin{aligned} m, n \text{ up to the order } h_{mn}(\xi)_0 \\ O, n \text{ „ „ „ „ } h_{0n}(\xi)_0 \\ O, O \text{ „ „ „ „ } h_{00}(\xi)_0 \end{aligned}$$

Let us assume now that we wish to go one step further. We then have the equations of motion:

$$\begin{aligned} \widetilde{A}^0_3 (\xi_0 + \xi_2 + \xi_4) + \widetilde{A}^0_5 (\xi_0 + \xi_2) + \widetilde{A}^0_7 (\xi_0) &= 0; \\ \widetilde{A}^m_4 (\xi_0 + \xi_2 + \xi_4) + \widetilde{A}^m_6 (\xi_0 + \xi_2) + \widetilde{A}^m_8 (\xi_0) &= 0. \end{aligned} \quad (10.6)$$

In $\widetilde{A}_3^0, \widetilde{A}_5^0, \widetilde{A}_4^m, \widetilde{A}_6^m$, the argument in (10.6) is different from that in (10.5) therefore they give contributions up to the eighth order. But to find $\widetilde{A}_0^m(\xi)$ we have to know

$$h_{mn}, h_{0m}, h_{00}, \quad (10.7)$$

all functions of ξ . Thus simply denoting:

$$Q^{a\beta} = -8\pi \left(T^{a\beta} - \frac{1}{2} g^{a\beta} T \right) \quad (10.8)$$

we have to solve the equations (omitting the a 's above the ξ, s):

$$\begin{aligned} R_{20}^{mn}(\xi + \xi) + R_{24}^{mn}(\xi) &= Q_{20}^{mn}(\xi + \xi) + Q_{24}^{mn}(\xi) \\ R_{30}^{0m}(\xi + \xi) + R_{50}^{0m}(\xi) &= Q_{30}^{0m}(\xi + \xi) + Q_{50}^{0m}(\xi) \end{aligned} \quad (10.9)$$

$$R_{20}^{00}(\xi + \xi + \xi) + R_{24}^{00}(\xi + \xi) + R_{44}^{00}(\xi) = Q_{20}^{00}(\xi + \xi + \xi) + Q_{24}^{00}(\xi + \xi) + Q_{44}^{00}(\xi)$$

This would seem to be an impossible task, since to solve the last equation (10.9) we would have to know ξ , which we wish to find by equation (10.6). However this is not so, because we have

$$R_{20}^{00} = -\frac{1}{2} h_{20,ss} = -4\pi \sum_{a=1}^p m^a \delta^a = Q_{20}^{00} \quad (10.10)$$

for *arbitrary* motion. Thus we may rewrite the last equation (10.9).

$$R_{40}^{00}(\xi + \xi) + R_{60}^{00}(\xi) = Q_{40}^{00}(\xi + \xi) + Q_{60}^{00}(\xi). \quad (10.11)$$

Collecting here the expressions of the sixth order we find $h_{60}(\xi)$. Similarly we can find h_{50m} and h_{4mn} . Thus we can push the approximation one step further.

We can now formulate the general theory.

In the development of $g_{a\beta}$; $R^{a\beta}$; $T^{a\beta}$, we took into account *arbitrary* motion. Under this assumption we developed, say

$$T^{mn} = T_{4}^{mn} + T_{6}^{mn} + T_{8}^{mn} + \dots \quad (10.12)$$

But, instead of *arbitrary* motion, let us put into the arguments *certain* motion developed into power series:

$$\xi = \xi_0 + \xi_2 + \xi_4 + \dots \quad (10.13)$$

and write, say,

$$T \left(\begin{smallmatrix} \xi & + & \xi & + & \xi \\ 2s & 0 & 2 & 4 \end{smallmatrix} \right). \quad (10.14)$$

If developed properly, the above expression gives us contributions of the order $2s + 2$ and $2s + 4$. Remembering this, let us assume that we have solved the equations of motion of the order $2r$:

$$\begin{aligned} \tilde{A}^0 \left(\begin{smallmatrix} \xi & + & \xi & + & \dots & + & \xi \\ 3 & 0 & 2 & 2r-4 & 5 & 0 & 2r-6 \end{smallmatrix} \right) + \dots + \tilde{A}^0 \left(\begin{smallmatrix} \xi \\ 2r-1 & 0 \end{smallmatrix} \right) = 0; \\ \tilde{A}^m \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 4 & 0 & 2r-4 & 6 & 0 & 2r-6 \end{smallmatrix} \right) + \dots + \tilde{A}^m \left(\begin{smallmatrix} \xi \\ 2r & 0 \end{smallmatrix} \right) = 0. \end{aligned} \quad (10.15)$$

This means, that we have solved the field equations in:

$$\begin{aligned} m, n \text{ up to the order } h_{mn} \\ 2r-4 \\ 0, m \text{ " " " " } h_{0m} \\ 2r-3 \\ 0, 0 \text{ " " " " } h_{00} \\ 2r-2 \end{aligned}$$

Now we wish to solve the equation of motion of order $2r + 2$:

$$\begin{aligned} \tilde{A}^0 \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 3 & 0 & 2r-2 & 2r+1 & 0 \end{smallmatrix} \right) = 0 \\ \tilde{A}^m \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 4 & 0 & 2r-2 & 2r+2 & 0 \end{smallmatrix} \right) = 0. \end{aligned} \quad (10.16)$$

The arguments in $\tilde{A}^0, \dots, \tilde{A}^0, \tilde{A}^m, \dots, \tilde{A}^m$ are different in (10.16) from those in (10.15). Therefore they give contributions of order $2r + 1$ and $2r + 2$. But to find $\tilde{A}^m \left(\begin{smallmatrix} \xi \\ 2r+2 & 0 \end{smallmatrix} \right)$ we have to know:

$$h_{mn}, h_{0m}, h_{00}.$$

We have, therefore, to solve the equations:

$$\begin{aligned} R^{mn} \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 2 & 0 & 2r-4 & 2r-2 & 0 \end{smallmatrix} \right) + \dots + R^{mn} \left(\begin{smallmatrix} \xi \\ 2 & 0 \end{smallmatrix} \right) = Q^{mn} \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 2 & 0 & 2r-4 & 2r-2 & 0 \end{smallmatrix} \right) + \dots + Q^{mn} \left(\begin{smallmatrix} \xi \\ 2r-2 & 0 \end{smallmatrix} \right); \\ R^{om} \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 3 & 0 & 2r-4 & 2r-1 & 0 \end{smallmatrix} \right) + \dots + R^{om} \left(\begin{smallmatrix} \xi \\ 3 & 0 \end{smallmatrix} \right) = Q^{om} \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 3 & 0 & 2r-4 & 2r-1 & 0 \end{smallmatrix} \right) + \dots + Q^{om} \left(\begin{smallmatrix} \xi \\ 2r-1 & 0 \end{smallmatrix} \right); \\ R^{00} \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 2 & 0 & 2r-2 & 2r & 0 \end{smallmatrix} \right) + \dots + R^{00} \left(\begin{smallmatrix} \xi \\ 2 & 0 \end{smallmatrix} \right) = Q^{00} \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 2 & 0 & 2r-2 & 2r & 0 \end{smallmatrix} \right) + \dots + Q^{00} \left(\begin{smallmatrix} \xi \\ 2r & 0 \end{smallmatrix} \right). \end{aligned} \quad (10.17)$$

Everywhere in these equations, with the exception of R^{00} and Q^{00} we substitute the motion already known. However $R^{00} = Q^{00}$ for an arbitrary motion. Then we can replace the last equations in (10.17) by

$$R^{00} \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 4 & 0 & 2r-4 & 2r & 0 \end{smallmatrix} \right) + \dots + R^{00} \left(\begin{smallmatrix} \xi \\ 4 & 0 \end{smallmatrix} \right) = Q^{00} \left(\begin{smallmatrix} \xi & + & \dots & + & \xi \\ 4 & 0 & 2r-4 & 2r & 0 \end{smallmatrix} \right) + \dots + Q^{00} \left(\begin{smallmatrix} \xi \\ 2r & 0 \end{smallmatrix} \right).$$

In R_{mn}^{2r-2} the expression h_{mn}^{2r-2} appears, for the first time. In R^{0m}_{2r-1} the expression h_{0m}^{2r-1} appears and finally in R^{00}_{2r} the expression h_{00}^{2r} appears. Collecting all the contributions of the highest order in these equations and putting them equal to zero we can find h_{mn}^{2r-2} , h_{0m}^{2r-1} , h_{00}^{2r} ; for h_{00}^{2r} the equation is purely a Poisson equation!

Looking back at equations (5.8), we see that if h_{0m}^{2r-1} is a solution of (10.17), then

$$\begin{aligned} h_{0m}^{2r-1} &= h_{0m}^{2r-1} + a_{0,m}^{2r-1} \\ h_{0mn}^{2r} &= h_{mn}^{2r} + a_{m,n}^{2r} + a_{n,m}^{2r}. \end{aligned} \quad (10.18)$$

is a solution also.

For example let us put $r = 2$; that is

$$\begin{aligned} h_{0m}^3 &= h_{0m}^3 + a_{0,m}^3 \\ h_{0mn}^4 &= h_{mn}^4 + a_{m,n}^4 + a_{n,m}^4 \end{aligned}$$

The choice of these functions can always be achieved by a coordinate-transformation from a coordinate-system in which the a 's equal zero. Then in the equations of motion of the 8-th order, the derivatives of these four functions will appear. Generally they can be so chosen as to annihilate the expressions of the 8-th order in the equations of motion. But it is difficult to judge whether such a coordinate-system would have any physical meaning. In any case, up to the post Newtonian approximation, the choice of the coordinate-system does not play any role as long as we stick to our approximation procedure by which its beginning is determined, that is the choice of h_{00}^2

and h_{mn}^2 .

APPENDIX A⁶

To distinguish between our δ -function and Dirac's δ -function we shall here denote the former by $\hat{\delta}$. Our aim is to give a "realistic" model, showing how to construct a sequence of $\hat{\delta}(\varepsilon)$ so that $\hat{\delta} = \lim_{\varepsilon \rightarrow 0} \hat{\delta}(\varepsilon)$ and such that for every ε :

$$\int_{-\infty}^{\infty} \hat{\delta}(\varepsilon) d_{(3)}x = 1; \quad \int \hat{\delta}(\varepsilon) r^p d_{(3)}x = 0; \quad p = 1, 2, \dots, k. \quad (A.1)$$

Such a model can be gained from a model $\delta(\varepsilon)$ of an ordinary Dirac δ -function satisfying the following conditions:

$$\delta(\varepsilon) = \delta(\varepsilon, r) = \frac{1}{\varepsilon^3} \Delta \left(\frac{r}{\varepsilon} \right) \quad (A.2)$$

⁶ This is an abbreviated and changed version of two papers written in collaboration with J. Plebański appearing in Bull. Acad. Pol. (1956 and 1957)

where $\Delta\left(\frac{r}{\varepsilon}\right)$ is such that

$$\frac{1}{4\pi} D^{(p)} = \int_0^\infty \Delta(z) z^{-p+2} dz; \quad p = 1, 2, \dots, k \quad (\text{A.3})$$

always exists, and

$$D^{(0)} = \int \delta(\varepsilon) d_{(3)}x = 4\pi \int_0^\infty z^2 \Delta(z) dz = 1. \quad (\text{A.4})$$

If $\delta(\varepsilon)$ does not have this property, it can be made to have it by multiplying it by $\left(\frac{r}{\varepsilon}\right)^k$ and renormalizing⁷. Thus with such δ 's we can form the model of our $\hat{\delta}$ function in the following way:

$$\hat{\delta}(\varepsilon, r) = \frac{1}{k!} \left(\frac{\partial}{\partial \varepsilon}\right)^k \left(\varepsilon^{k-3} \Delta\left(\frac{r}{\varepsilon}\right)\right) \quad (\text{A.5})$$

We have to show, that such a choice of δ , satisfies (A.1). To do so, let us start with the first equation (A.1):

$$\int \hat{\delta} d_{(3)}x = \frac{4\pi}{k!} \left(\frac{\partial}{\partial \varepsilon}\right)^k \varepsilon^k \int_0^\infty \Delta(z) z^2 dz = D^{(0)} = 1. \quad (\text{A.6})$$

Now as to the second equation (A.1):

$$\int \hat{\delta} r^{-p} d_{(3)}x = \frac{4\pi}{k!} \left(\frac{\partial}{\partial \varepsilon}\right)^k \varepsilon^{k-p} \int_0^\infty \Delta(z) z^{2-p} dz = \frac{1}{k!} \left(\frac{\partial}{\partial \varepsilon}\right)^k \varepsilon^{k-p} D^{(p)} = 0 \quad (\text{A.7})$$

for every integer p , if

$$1 \leq p \leq k.$$

Thus the $\hat{\delta}$'s defined by (A.5) satisfy (A.1).

This procedure can easily be generalized. We can introduce modified Dirac's functions $\hat{\hat{\delta}}$ from conditions

$$\int \hat{\hat{\delta}} r^{-p} d_{(3)}x = \omega_{(p)}; \quad p = 1, 2, \dots, k, \quad (\text{A.8})$$

⁷ e.g.: Let us take $\delta(\varepsilon) = (2\pi)^{-3/2} \varepsilon^{-3} e^{-1/2 \varepsilon^2 r^2}$, that is $\Delta(z) = (2\pi)^{-3/2} e^{-1/2 z^2}$. Such $\Delta(z)$ shall be changed into $\Delta(z) = (2\pi)^{-1} 2^{-\left(\frac{k+3}{2}\right)} \left[\Gamma\left(\frac{k+3}{2}\right)\right]^{-1} z^k e^{-1/2 z^2}$

where $\omega_{(p)}$ are arbitrarily prescribed numbers. Indeed the realistic $\hat{\delta}(\varepsilon)$ satisfying (A.8) in the limit $\varepsilon \rightarrow 0$ is the following:

$$\hat{\delta}(\varepsilon) = \sum_{s=0}^k (D^{(k-s)})^{-1} \frac{\omega_{(k-s)}}{s!} \left(\frac{\partial}{\partial \varepsilon} \right)^s \left[\varepsilon^{k-3} \Delta \left(\frac{r}{\varepsilon} \right) \right]. \quad (\text{A.9})$$

We find.

$$\begin{aligned} \int \hat{\delta}(\varepsilon) r^{-p} d_{(3)}x &= 4\pi \sum_{s=0}^k (D^{(k-s)})^{-1} \frac{\omega_{(k-s)}}{s!} \left(\frac{\partial}{\partial \varepsilon} \right)^s \varepsilon^{k-p} \int_0^\infty z^{-p+2} \Delta(z) dz \\ &= \sum_{s=0}^k (D^{(k-s)})^{-1} \frac{\omega_{(k-s)}}{s!} \left(\frac{\partial}{\partial \varepsilon} \right)^s \varepsilon^{k-p} D^{(p)}. \end{aligned} \quad (\text{A.10})$$

This is different from zero and finite for $\varepsilon \rightarrow 0$ only for $s = k - p$. We have:

$$\int \hat{\delta}(\varepsilon) r^{-p} d_{(3)}x = \omega_{(p)} + o(\varepsilon)$$

and for $\varepsilon \rightarrow 0$ we have (A.8).

We see that the use of Dirac's functions requires the prescription for the values of $\omega_{(p)}$'s. The one used in this paper is the most convenient for our purpose; it requires $\omega_{(p)} = 0$; $p = 1, 2, \dots, k$.

APPENDIX B

R_{00} up to the fourth approximation equals:

$$R_{00} = -\frac{1}{2} \varphi_{,ss} - \frac{1}{2} h_{00,ss} + h_{0s,0s} - \frac{3}{2} \varphi_{,00} + \frac{1}{2} \varphi_{,s} \varphi_{,s} - \frac{1}{2} \varphi \varphi_{,ss}. \quad (\text{B.1})$$

Therefore

$$\begin{aligned} R^{00} &= R_{00} (1 + 2h^{00}) = R_{00} (1 - 2\varphi) \\ R^{00} &= R^{00} \sqrt{-g} = (1 - \varphi) R^{00} = R_{00} (1 - 3\varphi) \end{aligned} \quad (\text{B.2})$$

In (B.2) exceptionally $g = |g_{\alpha\beta}|$. We have:

$$R^{00} = -\frac{1}{2} \varphi_{,ss} + \varphi \varphi_{,ss} + \frac{1}{2} \varphi_{,s} \varphi_{,s} + \frac{1}{2} \varphi_{,00} - \frac{1}{2} h_{00,ss}. \quad (\text{B.3})$$

The right hand side of our gravitational equations

$$R^{a\beta} = -8\pi \left(T^{a\beta} - \frac{1}{2} g^{a\beta} T \right) \quad (\text{B.4})$$

is up to the fourth order for the "zero-zero" component:

$$-8\pi \left[T_{2}^{00} - \frac{1}{2} (\eta^{00} - \varphi) T_{2}^{00} (\eta^{00} + \varphi) + T_{4}^{00} - \frac{1}{2} T_{4}^{00} + \frac{1}{2} T_{4}^{ss} \right] \quad (B.5)$$

$$= -4\pi (T_{2}^{00} + T_{4}^{00} + T_{4}^{ss})$$

Because

$$\frac{1}{4} m = \frac{1}{2} \frac{1}{4} m \dot{\xi}^s \dot{\xi}^s + \frac{1}{r} \frac{1}{2} m m$$

$$\frac{2}{4} m = \frac{1}{2} \frac{2}{2} m \dot{\eta}^s \dot{\eta}^s + \frac{1}{r} \frac{1}{2} m m$$

we have for the right hand side:

$$-4\pi \left(\frac{1}{2} \frac{1}{2} m \delta + \frac{2}{2} \frac{2}{2} m \delta + \frac{3}{2} \frac{1}{2} m \dot{\xi}^s \dot{\xi}^s + \frac{3}{2} \frac{2}{2} m \dot{\eta}^s \dot{\eta}^s + \frac{1}{r} \frac{1}{2} m m \delta + \frac{1}{r} \frac{2}{2} m m \delta \right) \quad (B.6)$$

Now on both sides we shall take only expressions α) of the order four, β) those that give a finite contribution to $\widetilde{h_{00,m}^{(4)}}$. Thus this relevant part of $h_{00}^{(4)}$ satisfies the equation:

$$h_{00,ss}^{(4)} = 2f g_{ss} + 2g g_{ss} + g_{,00} + g_{,ss} g_{,ss} + 8\pi \left(\frac{3}{2} \frac{2}{2} m \dot{\eta}^s \dot{\eta}^s \delta + \frac{1}{r} \frac{2}{2} m m \delta \right) \quad (B.7)$$

Because

$$g_{ss} = 8\pi \frac{2}{2} m \delta; \quad f = -2 \frac{1}{m} \left(\frac{1}{r} \right)^{-1} \quad (B.8)$$

we have for the contribution of C to h_4 coming from the first two expressions:

$$C_{ss}^{(4)} = -4\pi \frac{2}{\delta} a. \quad (B.9)$$

Generally the solution of

$$C_{ss}^{(4)} = -4\pi \frac{2}{\delta} a \quad (B.10)$$

is

$$C = \frac{2}{a} \left(\frac{2}{r} \right)^{-1} \quad (B.11)$$

where

$$\frac{2}{a} = \int \frac{2}{\Omega(s)} a \delta d_{(3)} x. \quad (B.12)$$

Thus in our case:

$$C = \frac{{}^4 m m}{{}_2 r r} \quad (B.13)$$

Therefore:

$$h_{00,ss} = g_{00} + g_{ss} g_{ss} + 8\pi \left(\frac{3}{2} m \eta^s \eta_s \delta - \frac{{}^1 m m}{{}_2 r} \delta \right) \quad (B.14)$$

Finally therefore, we have:

$$h_{00} \rightarrow -m {}_2 r_{,00} + 2 \binom{2}{m} \binom{2}{r} - 3 m \eta^s \eta_s \binom{2}{r}^{-1} + 2 m m \binom{2}{r r}^{-1} \quad (B.15)$$

Let us now generalize h_{00} for *three* particles, again only looking for expressions which give a contribution to $\widetilde{h_{00,m}}$. The only non trivial expression of this kind are those proportional to $m m$.

We denote by ${}^{(a b)} r$, the "distance" between the a and b particle⁸:

$$\binom{(ab)}{r}^2 = \binom{a}{\xi^s - \xi^s} \binom{b}{\xi^s - \xi^s} \quad (B.16)$$

and we ask: what are the contributions to (B.4) coming from the third particle and proportional to $m m$? We have now:

$$\varphi = h_{00} = f + g + k \quad (B.17)$$

$$f = -2 m \binom{1}{r}^{-1}; \quad g = -2 m \binom{2}{r}^{-1}; \quad k = -2 m \binom{3}{r}^{-1}.$$

Then the additional expressions for which we look in (B.3) are:

$$\frac{1}{2} (g k)_{,ss} + \frac{1}{2} g k_{,ss} + \frac{1}{2} k g_{,ss} - \frac{1}{2} S_{00,ss} \quad (B.18)$$

where S_{00} denote the additional expression in h_{00} . The additional expressions in (B.5) because of (B.6) are:

$$-4\pi m m \binom{2}{r} \binom{(23)}{r}^{-1} \binom{2}{\delta + \delta}^{-1}. \quad (B.19)$$

Therefore the additional expression to $h_{00,ss}$ is:

$$S_{00,ss} \sim 8\pi m m \binom{3}{r} \binom{(23)}{r}^{-1} \binom{2}{\delta + \delta}^{-1} + (g k)_{,ss} + g k_{,ss} + g k_{,ss} \quad (B.20)$$

From this we find

$$S_{00} = 2 m m \left[\binom{(23)}{r}^{-1} \binom{3}{r}^{-1} + \binom{(23)}{r}^{-1} \binom{2}{r}^{-1} + 2 \binom{23}{r r}^{-1} \right]. \quad (B.21)$$

⁸ Previously ${}^{(1 2)} r = r$

Therefore, as we see, S_{00} does not depend explicitly on ξ . Therefore:

$$\widetilde{S}_{00,m} = \widetilde{S}_{00,m} \quad (\text{B.22})$$

and

$$\frac{1}{2} \widetilde{S}_{00} - \frac{1}{4} \widetilde{g}^k = \frac{2}{m} \frac{3}{m} \left[\left(\begin{smallmatrix} (23)(31) \\ r \ r \end{smallmatrix} \right)^{-1} + \left(\begin{smallmatrix} (23)(21) \\ r \ r \end{smallmatrix} \right)^{-1} + \left(\begin{smallmatrix} (12)(13) \\ r \ r \end{smallmatrix} \right)^{-1} \right]. \quad (\text{B.23})$$

The last very simple question with which we shall deal here is the change from h_{0m} to h'_{0m} :

$$h'_{0m} = h_{0m} + a_{0,m}. \quad (\text{B.24})$$

Putting this into (B.1), we have

$$h'_{00} = h_{00} + 2a_{0,0} \quad (\text{B.25})$$

which is identical with (9.4).

КРАТКОЕ СОДЕРЖАНИЕ

Л. Инфельд, Уравнения движения в общей теории относительности и принцип наименьшего действия.

В настоящей работе обсуждён способ формулирования уравнений движения для p тел, принимая: 1) уравнения поля общей теории относительности, 2) факт, что тензор энергии скорости отличается от нуля единственно вдоль p мировой линии известной формы, иначе говоря, что тензор этот линейно зависит от δ -функции Дирака.

Показано, что на основании этих простых предпосылок можно получить надлежащий лагранжиан, избегая при этом обременительных исчислений. Лагранжиан этот ведет к вьеньютоновским уравнениям движения.

Выяснена тоже связь между полученными уравнениями движения и уравнениями „геодезических линий“.

REFERENCES

- Einstein A., Grommer J., *Sitz. Berl. Akad. Wiss.* 2, (1927).
 Einstein A., Infeld L., and Hoffmann B., *Ann. Math.* **39**, 66 (1938).
 Einstein A., Infeld L., *Ann. Math., Can. J. Math.*, **1**, 209 (1949).
 Fichtenholz I. G. *Journ. Exp. Teor. Fis.* **27**, 563 (1954).
 Fock W. A., *J. Phys. (Moscow)* **1**, 81 (1939).
 Infeld L., *Phys. Rev.*, **53**, 836 (1938), *Can. J. Math.*, **5**, 17 (1953), *Acta Phys. Pol.* **13**, 205 (1954).
 Infeld L. and Plebański J., *Bull. Acad. Pol.* (in print).
 Papapetrou A., *Proc. Phys. Soc. (London)*, **64**, 57 (1951).
 Petrova, *Izh. Eksper. Teor. Fiz.* **19**, 989 (1949).
 Rameswararao B., *Thesis, of the Banaras Hindu University* (1955).
 Scheidegger, A. E. *Rev. Mod. Phys.* **25**, 451 (1953).
 Tisseyre R., *Acta Physica Polon.* **13**, 47 (1954).
 Tulczyjew W., *Bull. Acad. Pol. Cl. III* **5**, 279 (1957).

ON THE FREE-ELECTRON THEORY OF ABSORPTION SPECTRA OF SOME LINEAR CONJUGATED SYSTEMS

BY S. OLSZEWSKI

Electrode Processes Laboratory, Institute of Physical Chemistry, Polish Academy, of Sciences, Warsaw

(Received December 5, 1956)

The AS FEMO method was applied in this work to calculate the absorption maxima in the spectrum of symmetrical cyanine dyes and polyenes. The results were discussed in comparison with experimental data and these of the simple free-electron model.

In the case of cyanines considerable improvement of the theoretical values is obtained by introducing the resonance barrier in the FE model. The height of this barrier is determined by the stabilization energies of the cyanine nuclei. The effect of the energy states obtained by the introduction of the barrier to the simple FE model is discussed in detail and the correspondence between the results of resonance theory for cyanines and those of a FE model with a barrier was indicated.

§ 1. Some Results of the Application of the AS FEMO Method

The recently developed AS FEMO (antysymmetrized free-electron molecular orbitals) method (Olszewski 1955) allows improved calculations to be made on the absorption spectra of non-branched bond systems of conjugated molecules in the free-electron scheme. In this method even long chains may easily be investigated.

Let us assume, as usual, that in the free-electron model, for the above molecules, the energies corresponding to the electronic eigenstates are:

$$E = E_n = \frac{h^2 n^2}{8mL^2} \quad (1)$$

where n is an integer,
and the eigenfunctions are:

$$\psi_n = (2/L)^{1/2} \sin \frac{n\pi}{L} s, \quad (2)$$

where L is the free-electron path.

This may be called a first approximation solution to the problem. (The interaction potential of electrons with the core is taken as equal to zero). In the second approxi-

mation (which corresponds to the AS MO treatment) we built the eigenfunctions from the antisymmetrized products of molecular spin-orbitals, the mutual electronic interaction operator being added to the Hamiltonian. We then can use the formula given below (Roothaan 1951) for the energy transition from the occupied ground state level i to the excited state level a :

$$E(^{1,3}\Phi_{ia}) - E(^1\Phi_0) = H_a + \sum_j (2J_{ja} - K_{ja}) - H_i - \sum_j (2J_{ij} - K_{ij}) + (J_{ia} - K_{ia}) \pm K_{ia}, \quad (3)$$

where the summation runs over all orbitals of the ground state; i.e. for an even number of π -electrons the summation is taken over $N/2$. The signs $+$ and $-$ correspond to the singlet and triplet states, respectively.

For the elements appearing in (3) in the case of the non-branched free-electron path, the AS FEMO method gives the following expressions:

$$H_n = E_n$$

$$J_{nm} = \frac{e^2}{\pi L} \left\{ \pi + \frac{m^2}{(n^2 - m^2)n} \text{Si}(2n\pi) - \frac{n^2}{(n^2 - m^2)m} \text{Si}(2m\pi) \right\},$$

$$J_{nn} = \frac{e^2}{\pi L} \left\{ \pi - \frac{3}{2} \frac{\text{Si}(2n\pi)}{n} \right\} = K_{nn}, \quad (4)$$

$$K_{nm} = \frac{e^2}{\pi L} \left\{ \text{Si}[(n+m)\pi] \left(\frac{1}{n+m} - \frac{n+m}{nm} \right) + \text{Si}[(n-m)\pi] \left(\frac{1}{n-m} + \frac{n-m}{nm} \right) \right\},$$

where n and m may obviously be exchanged.

The scheme given above has been applied to the polyenes and symmetrical cyanines with various N . The wave lengths corresponding to the lowest electronic transitions ($N/2 \rightarrow N/2 + 1$) are calculated¹ in Tables I & II.

For L the $(N+1)d$ value was taken for polyenes, and for symmetrical cyanines — Nd , where $d = 1.40 \text{ \AA}$ is the mean length of the C — C bond in the chains.

Theoretical results compared with the experimental data for polyenes show that the free-electron model cannot be applied to the treatment of these molecules. To λ_{max} there correspond much shorter waves than those obtained by the theory, and the absolute value of the difference between $\lambda_{\text{AS FEMO}}$ and λ_{exp} increases with increasing chain length. This situation is caused mainly by the well-known fact, that the π -electrons in polyenes do not form a complete free-electron gas, but are localized on the double

¹ In the case of polyenes with $N = 2$ (ethylene) and $N = 4$ (butadiene) the exact data are given by Olszewski (1955).

TABLE I
The locations of absorption maxima for polyenes

Polyen	N	λ (Å) AS FEMO		λ (Å) simple FE model	λ (Å) exp.
		singul.	triplet		
Hexatrien	6	3170	4590	4530	2600
Oktatetraen	8	3930	5800	5820	3020
Axerophthene	10	4750	7150	7110	3460
Anhydro-Vitamin A	12	5450	8220	8400	3690
Dimethylpolyen	14	6420	9930	9700	3995*
Dimethylpolyen	16	7110	10990	10990	4090*
Dimethylpolyen	18	7920	12310	12280	4160*
Dimethylpolyen	20	8700	13570	13570	4230*
β -Carotene	22	9500	14890	14870	4510
Dehydro β -Carotene	24	10290	16160	16160	4750
	26	11090	17470	17450	
	28	11870	18740	18740	
Dehydrolycopene	30	12670	20060	20040	5040

The quantities with an asterisk are taken from the work: Nayler, Witing (1954), the remaining other experimental data from Kuhn, (1949).

bonds of the molecule. Nevertheless, in the AS FEMO scheme for λ_{\max} , we generally get values closer to the experimental results than in the simple free-electron model. This is mostly true for the singlet positions, which lie at about half way between the experimental value and that taken from the simple model.

Kuhn (1949) has obtained good agreement with experiment for polyenes by introducing in the simple scheme a sinusoidal potential along the molecular core. In this picture the potential wells corresponded to π -electrons placed on the double bonds. The present results give evidence to the fact that the potential amplitude intro-

TABLE II
The theoretical locations of absorption maxima for symmetrical cyanines

N	$\lambda_{\text{AS FEMO}}$ sing. & tripl. arith. mean	λ_{FE} simple model	N	$\lambda_{\text{AS FEMO}}$ sing. & tripl. arithm. mean	λ_{FE} simple model
6	2900	3320	14	7180	8450
8	3905	4595	16	8085	9730
10	4980	5880	18	9140	11020
12	5895	7160	20	10170	12310

duced by Kuhn should be smaller, i.e. the localization of electrons is not as strong as that previously considered².

Agreement between the theoretical calculations and the experimental data for the symmetrical cyanines constitutes another problem. This will be discussed in more detail in the following sections. It should now be emphasized that in comparison with the simple model the wavelengths of the absorption maxima are now shifted, as in the case of polyenes, in the direction of the shorter waves. The vinylene shift is less (approx. 1050 Å instead of 1300 Å) than for a simple model, its value, however, remains constant even for large N . Such behaviour is seen from the experimental data for the spectrum of symmetrical cyanines, and agreement between the theoretical and experimental numerical shift values is now apparent.

The results of the application of the AS FEMO scheme may be refined by including the configurational interaction. It follows, however, from the calculations for ethylene and butadiene (see footnote 1) that the effect of this interaction is rather small. Nevertheless, it should be expected that it may increase with the chain length, i.e. with an increase in the number of π -electrons. At the same time the terms occurring in the $E(1^3\Phi_a) - E(1\Phi_0)$ which consist of Coulomb and exchange integrals will begin to play a more important role.

It is characteristic for the calculations that the triplet levels fall almost exactly at the same place as the energy levels of a simple free-electron model. There seems to be no explanation for this fact at present.

§ 2. The Concept of the Resonance Barrier

In the symmetrical cyanine cation two principal elements may be distinguished:

1) — a chain of CH groups connected in turn by double and single bonds. The nitrogen atoms are placed at both of its ends.

2) — a "nucleus" composed of the system of those atoms at one of the ends of the molecule, which do not form a chain of conjugated double bonds, and of those atoms of the chain which are directly bound with this system. As a rule from 1 the nitrogen atom and often one or more carbon atoms will belong to the nucleus. Symmetrical cyanine evidently has the same nuclei at both ends.

There still exists a radical at the nitrogen atom. We exclude it from the system which we understand as the "nucleus".

² However, the free-electron model can be applied with quite good results to polyenes with a small number of π -electrons. Then the π -electron density in the ground state $\left(2\sum_n^{N/2}\psi_n^2\right)$ clearly corresponds to the concentration of these electrons in double bond regions (see Fig. a); for hexatriene see Platt (1954). With increasing chain length the maxima become flattened, we obtain a uniform density distribution, which corresponds to symmetrical cyanines. At the same time it is interesting to note that the "experimental" potential amplitude given by Kuhn also decreases with the chain length.

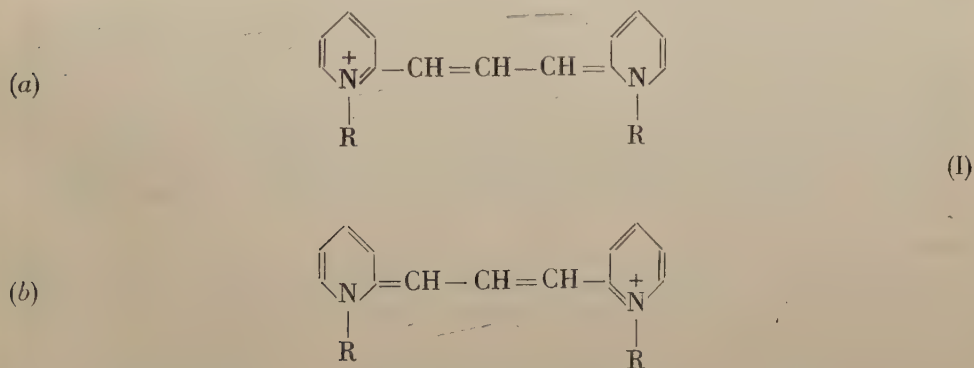
The symmetrical cyanines are characterized by the equal probability that in the conjugated chain between any two atoms there exists a double or single bond. It follows that the charge is uniformly distributed along the atomic core, i.e. there is no collection of charge in the space corresponding to the bond of the pair of neighbouring atoms when the charge on the other bond is decreased.

We would not, however, venture to say that the charge distribution is the same in all parts of the core. The end groups interact with the chain and the effect of this interaction will be, as we shall see, particularly strong in the regions of the conjugated system which are included in the nuclei. Outside the nuclei we shall assume, further, that the chain is approximately "unperturbed" and the charge is distributed uniformly.

The electrons of the conjugated system in part of the cyanine nucleus are components of the electronic structure of the nucleus, as well the electronic structure of the chain. The nucleus forms a molecular system generally quite different and isolated from the rest of the chain. Thus, it often happens that the part of the chain corresponding to the nucleus plays a more important role in the structure of the nucleus as a whole than in the structure of the conjugated system.

Now the electronic structure of the nucleus tends to take on the lowest energy state and the actual charge distribution will no longer be represented by the chemical formula of the nucleus, since this tendency of the electron assembly to form the most stable system has to be taken into account. In this way the nuclei act as if they forced the electrons of the conjugated chain to take part in their most convenient (i.e. from the energetical point of view) structure.

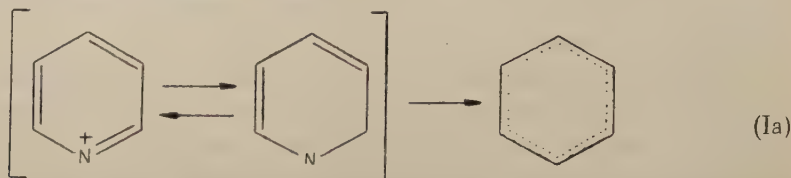
Let us illustrate our discussion by the example:



In the former theory the actual charge distribution in the molecule is the intermediate value between (a) and (b). In the conjugated part of the nucleus this value lies between the state when both π -electrons are closely linked with the nitrogen atom and that, when they are distributed along the N — C bond forming a double bond. In the rest of the nucleus the charge distribution for both cases (a) and (b) remains unchanged; there are two conjugated double bonds. Together with a pair of electrons

from the conjugated part they form a system very similar to that of three double bonds in the hexagonal ring, i.e. the Kekulé structure of benzene.

Such a system, as we know, is very unstable and reveals a strong tendency to admit lower energy corresponding to a uniform charge delocalization along the hexagonal core. Owing to their share in the electronic structure of the ring, the electrons from the chain tend to a state of much lower energy than that which would correspond to their participation in the conjugated system only.



The above considerations have their analogue in the theory of the resonance structures. In (I) two extreme resonance structures of a dye are shown. However, the actual state of the molecule cannot be obtained only from the resonance of the extreme structures. The interaction of the intermediate structures, corresponding to the position of positive charge at the C atoms in the chain, must also be taken into account. When the positive charge is placed on any of the nitrogen atoms we have a system of three double bonds conjugated in a ring. The stable benzene state is then more easily attained than in the case where there is no + charge on the N atom, i.e. when both electrons are in the closest neighbourhood of N. Therefore the limiting structures have lower energies than the intermediate structures. (See discussion of Brooker (1942)). We shall further discuss this analogy later on.

Let us now extend our treatment to the free-electron theory. It is well-known that the symmetrical cyanines seem to be best investigated by means of a free-electron model, because of the equivalence of the bonds in the chain.³

In the box of the free-electron model we distinguish a part corresponding to the chain between the two nuclei and parts corresponding to the chain included in the nuclei themselves. According to the above discussion, in the first case we have, following Kuhn, the periodic potential, which for simplicity, we approximate by a constant value. Two deep potential wells caused by the attraction of the moving (π -) electrons by the end groups correspond to the other parts. Thus a potential barrier appears (see Fig. 1).

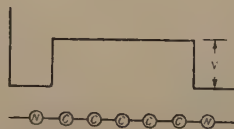


Fig. 1. The free-electron model with resonance barrier

³ The decrease in the potential at the ends, which is caused by the presence of the nitrogen atoms alone, these atoms being more electronegative than the carbon atoms, is not considered here. This effect is rather small in comparison to the resonance barrier height, and *does not* select the cyanines.

Hereafter this will be called the *resonance barrier*, for, as we shall see later, there exists an analogy between our barrier and the intermediate structures in the resonance theory.

We assume for simplicity that the barrier has a rectangular shape and that the ends of the barrier lie at the middle of the C — N bonds. This seems justified in most cases. The length of the free-electron path is equal to that in the problem without a barrier.

Thus we have a quantization problem of free-electron motion in a box, where the potential at the ends is equal to zero, and a potential barrier of a constant height V is stretched over the carbon atoms.

Assuming that at the particular sections of the box the solutions are given by the free-electron eigenfunctions⁴ and taking the advantage of the condition of continuity of the functions and their derivatives, we get for the eigenenergies E_x of the barrier problem the following equation⁵:

$$\begin{aligned} & \cos \left(\frac{3\pi}{Z+1} x \right) \sin \left[\pi x \frac{Z-2}{Z+1} \left(1 - \frac{a^2}{x^2} \right)^{1/2} \right] + \\ & + \sin \left(\frac{3\pi}{Z+1} x \right) \cos \left[\pi x \frac{Z-2}{Z+1} \left(1 - \frac{a^2}{x^2} \right)^{1/2} \right] \cdot \left(1 - \frac{a^2}{x^2} \right)^{1/2} + \\ & + \sin^2 \left(\frac{3}{2} \cdot \frac{\pi}{Z+1} x \right) \sin \left[\pi x \frac{Z-2}{Z+1} \left(1 - \frac{a^2}{x^2} \right)^{1/2} \right] \frac{a^2}{x^2} = 0, \end{aligned} \quad (5)$$

$Z (= N + 1)$ is the number of atoms in the chain;

$$V = \frac{h^2 a^2}{8mL^2}; \quad E_x = \frac{h^2 x^2}{8mL^2} \quad (6)$$

It is now essential to find V . The value of V must be closely connected with the ability of the cyanine end groups to form stable systems. The measure of this ability is the energy difference of the unstable system, i.e. the nucleus without π -electrons of the chain as well as the π -electron pair in C — N region, and the energy of a stable system, i.e. the nucleus as a whole. We shall call this energy difference the stabilization energy. It may also be obtained in a rough approximation on the basis of the free-electron model.

The energy of a π -electron pair in the C — N region is greater than that for a common double bond. We assume that it corresponds to the energy of a pair of electrons in a box two C — C bonds long (instead of two bonds for a common double bond system as in the case of ethylene).

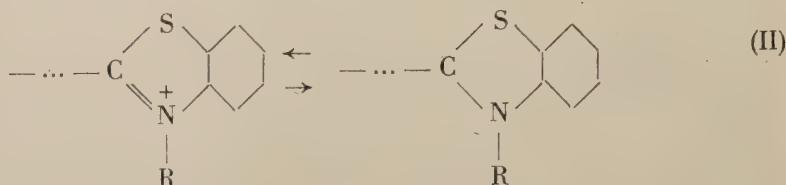
For (I, Ia) we divide the ring forming the nucleus into two segments: 1° — the free-electron path for C — N; 2° — the path in a system of two conjugated double bonds. Thus the stabilization energy is given by

$$E_{stab.} = E_{1^\circ} + E_{2^\circ} - E_{ring} \quad (7)$$

⁴ The electronic interaction is neglected here.

⁵ An equivalent equation was considered by Kuhn (1951) in his discussion of the effect of the electronegativity differences of the core atoms. However the present form of equation is better suited to our further discussion of the energy states.

For cyanine with a nucleus of the type



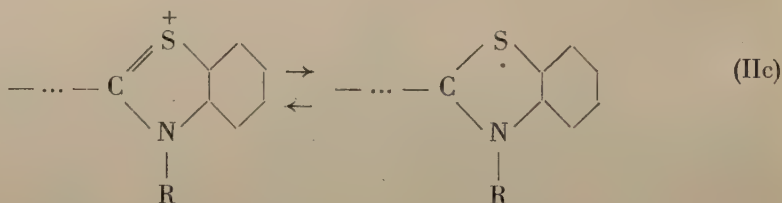
the stabilization energy corresponds approximately to the difference of energies for a ring and a C—N segment



and the system



where the free path is also lengthened by one bond on the last atom, i. e. we get a system of styrene type. In addition, besides the (II)-scheme an alternative system is also possible:



therefore the stabilization energy for this nucleus ought to be calculated twice.

The singular stabilization energy of the above type is realized approximately by:



where it is assumed that the interaction of the conjugated chain with the ring through the C—C bond is weak and that only the π -electron pair from C—N region takes part in the stabilization of the nucleus. In fact, we see that for the same number of π -electrons ($N = 10$) in the conjugated chain there is a large difference in the wavelengths of the absorption maxima in cyanines (II) and (III): for (II), where there is a higher barrier, this length is greater.

When in cyanine nucleus (in the non-conjugated part) there are no additional double bonds (e.g. thiazoline system) then the end groups are weakly stabilized. The barrier is low and caused mainly by the presence of the N atoms in the chain.

The stabilization energies for various cyanine nuclei are given in Table VI⁶. Later on they will be treated as V in the barrier problem.

§ 3. The Effect of the Resonance Barrier on the Energy States

We shall discuss the effect of the barrier of a variable height V (or a^2) and of different lengths corresponding to $Z = 5, 7, 9$. From the numerical solutions of (5) it follows that x increases proportionally to V (see Fig. 2). This linear dependence is different for a particular n and Z , and is well satisfied for not too high V . For larger V the straight lines are slightly bent, forming approximately arcs of parabolas with a^2

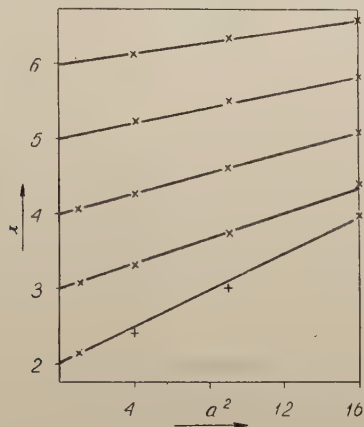


Fig. 2. The numerical solutions of Eq. (5) for the case $Z = 5$ ($N = 6$) and $a^2 \leq 16$. The origins of the straight lines correspond to the case when the barrier does not exist ($a^2 = 0$). We see directly from (5) that x is then equal to n from (1), i.e. we obtain the states of a simple free-electron model. The \times are associated with the exact solutions

as the axis. In addition it appears that at a fixed Z (a constant barrier length) the directional coefficients of the lines corresponding to increasing n 's — decreases. Also the parabolas tend to meet. (For numerical results see Tables III and IV).

⁶ For the nuclei given in Table VI, which have not been considered above, we calculate the approximate stabilization energies from the expressions:

for number 4 — $2[E_{C-N} + E_{\text{ethyl. syst.}} - E_{\text{butad. syst.}}]$

for number 5 — $(E_{\text{styrene syst.}} + E_{C-N} - E_{\text{napht. syst.}} + E_{\text{benz. syst.}} + E_{C-N} - E_{\text{styr. syst.}})$

The values for the eigenenergies of some systems were taken from Scherr (1953). For energy calculations the mean length of the bond was taken equal to 1.39 Å (as for ring systems).

TABLE III

The — m — directional coefficients for $x = ma^2 + n$ lines corresponding to the approximate solution of eq. (5) for various n and Z

The linearity is preserved well enough for $Z = 5$ up to $a^2 = 16$; for $Z = 7$ up to $a^2 = 25$; for $Z = 9$ up to $a^2 = 36$. Further on the lines for greater a^2 change into parabolas.

$\begin{smallmatrix} n \\ Z \end{smallmatrix}$	2	3	4	5	6
5	0.1205	0.0831	0.0683	0.0526	0.0383
7		0.0917	0.0715	0.0620	0.0534
9		0.109	0.0769	0.0631	0.0559

TABLE IV

The — b — and — c — coefficients in parabolic eq. $a^2 = bx^2 + c$ ($x > 0$) corresponding to the approximate solution of eq. (5) for greater a^2 . (See explanatory notes to Table III).

b	$\begin{smallmatrix} n \\ Z \end{smallmatrix}$	3	4	5	6
	5	1.217	1.767		
	7		1.203	1.424	
	9			1.142	1.240
c	$\begin{smallmatrix} n \\ Z \end{smallmatrix}$	3	4	5	6
	5	—7.975	—31.12		
	7		—16.70	—37.27	
	9			—24.79	—43.75

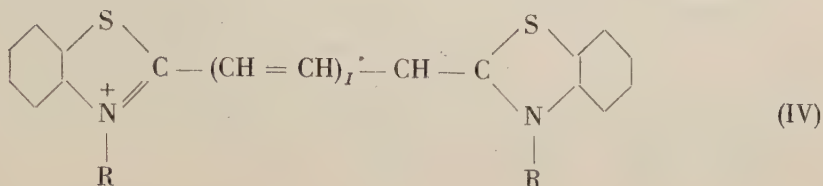
Thus, the lower the given quantum state, the greater, in general, the increase in energy for the given barrier height as compared to the corresponding eigenenergy of the nonbarrier problem.

On the other hand it follows from the calculations that at a fixed height of the resonance barrier, and for a given n, x (and subsequently E_x) increases with the chain length. This may be treated as a parallelism to Moffitt's conclusion from his analysis of absorption spectra of symmetrical cyanines on the basis of resonance structures theory (1950). With an increase in the number of $\text{CH} = \text{CH}$ groups in the chain, the number of intermediate structures ("bridge" structures) is increasing, and the difference between the given state for the case when these structures are not considered and the state obtained by introduction of all the structures also increases.

For a constant barrier height and for very large Z , eq. (5) gives solutions corresponding to results obtained by Kuhn for the constant potential over the entire box of a free-electron model. This potential causes a shift of all the energy levels by a constant value V . Thus for very long chains the effect of the resonance barrier with a fixed height on the spectrum ceases to play essential role.

In general, if the height of a resonance barrier is known, the wavelength of the appropriate absorption band may be calculated (Table VI), and vice versa.

We shall evaluate for the latter case the height of the resonance barrier for various chain lengths starting with the experimental data of the spectrum of cyanine:



Thus we obtain:

I	λ (Å) exp.	barrier height (eV)	λ (Å) simpl. model
0	4230	11.6	3320
1	5575	11.3	4595
2	6500	10.1	5880

and the theoretical barrier height (the stabilization energy) lies above 9.8 eV. Hence it follows that with an increase in chain length, the barrier corresponding to the particular (IV) homologues subsequently decreases. This may be explained perhaps by the growth of the self-stabilization of a cyanine chain with an increase in the length of this chain. The end-group effect (as a relative stabilization effect) is smaller for the longer systems than for the shorter ones.

We may regard this result as being experimentally proved by the fact that the ability to bind a proton to the cyanine (IV) chain increases with an increase in the number of $\text{CH} = \text{CH}$ groups (Brooker 1942).

When the proton is joined to the chain, it spoils the delocalization of π -electrons and we get a system partly similar to the model for polyenes. As mentioned before a periodic potential corresponds to such a model (Kuhn). Then the mean value of the potential is close to zero, so the potential barrier is destroyed here. We could expect that the facility with which the barrier is destroyed is connected with the barrier height.

§ 4. Conclusions

When the free-electron model is adapted to the treatment of absorption spectra of symmetrical cyanines, a question may arise as to what extent this model fulfils its purpose.

The old simple FE model, developed for cyanines by Kuhn, has given good agreement with experiment only for a few molecules of this kind. For the majority of cyanines the λ_{max} experimental values were less or greater than expected by the model; for the latter case the differences used to be very large. The model did not explain the essential question: *Why* the conjugated systems in cyanines of the *same* number of π -electrons often gave big differences in the absorption wave-lengths?

Kuhn (1948, 1949) tried to explain the absorption in the longer wave-length region by assuming an additional lengthening of the free-electron path, but he did not give any theoretical estimate of it.⁷ For the shorter wavelengths one should have then assumed a shorter path than that for a free-electron in the molecule. Now, by a more precise treatment of the AS FEMO method *all* the experimental values for λ_{\max} lie in the region of wavelengths longer than those predicted from the model. However, in the case of molecules for which the physical conditions of a pure free-electron model are best satisfied, λ_{\max} should be in quite good agreement with the AS FEMO values.

In fact, for molecules (Simpson cyanine dyes)



in which there is no stabilization in the end-groups, we obtain:

TABLE V
The Simpson cyanine-dyes

j	N	λ AS FEMO	λ Experim.
0	6	2900	3090
1	8	3905	4090
2	10	4980	5110

When we introduce the resonance barrier into the simple free-electron model, we can roughly, at least, overcome the difficulties which arise in the interpretation of the spectra by a non-barrier model. As follows from the solutions in Sec. 3, the presence of a high barrier raises the ground state more than an excited state. Thus the difference between these states decreases and *we get absorption at longer wave lengths than for the non-barrier model*, as is observed. This change of the position of states is determined, as we have seen before, by the barrier height, and therefore it is a function of the stabilization energy of the cyanine nuclei.

In our picture the evident shift towards the long waves appears, as a rule, for greater barrier heights. We very often deal with cyanines having such high barriers. For lower barriers the results are generally not much different from those obtained on the basis of a simple model. We even get absorption in wavelengths shorter than

⁷ In addition to Kuhn, the following authors, on the basis of empirical data, also took into account the effect of the end nuclei: Nikitine (1951), Basu (1954a) fitted the length of the free-electron path; Fukui et al. (1953) fitted the periodic potential in the box, Simpson (1948) fitted both the C — C bond distance and the path length. The non-empirical treatment of Basu (1954b) is rather doubtful in view of the asymmetry of the potential in the box and very good results seem to be obtained for an exceptional case only; see, however, footnote 9.

predicted by the model, the deviations, however, are small. Thus we may conclude that, in general, for barriers that are not high the λ_{\max} are close to the λ from the non-barrier model, as is actually the case⁸. (Table VI).

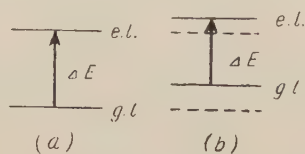


Fig. 3. The shift in the energy levels of the ground state (g.l.) and the excited state (e.l.) caused by the existence of a high resonance barrier.

(a) free-electron model without a barrier

(b) a barrier model

The change in the energy corresponding to the absorption (ΔE) is significantly greater in (a) than in (b)

We shall now discuss briefly the vinylene shift problem. Let us calculate the wavelengths of absorption maxima for a fixed barrier height in the case of several subsequent homologues produced by the increase in the number of vinylene groups of the chain. We find that the change in λ_{\max} corresponding to the two subsequent homologues (the vinylene shift), particularly at the high barriers, is much greater when, for instance $N = 6 \rightarrow N = 8$, than for $N = 8 \rightarrow N = 10$. For the transitions corresponding to still greater N this shift approaches that given by the non-barrier free-electron model.

For "3" cyanine, for example, the first shift calculated from the barrier model is greater by 150 Å than the next one. (Table VI). The simple model does not show any difference in the shift here.

Thus we obtain the theoretical explanation of the well-known experimental fact that the first vinylene shift for the given series of cyanine is greater than the subsequent values of this shift.⁹

⁸ From the numerical point of view the differences between the results of the barrier model and the experimental data sometimes still remain serious. On the other hand, the roughness of our method is apparent. To improve it the effect of the radicals on the ends and the difference in the electronegativities between C and N should be taken into consideration in addition to the theoretical refinement of the calculations.

⁹ Such behaviour of the vinylene shift was first communicated by Brooker (1942). He listed cyanine "3" as a typical compound (Table VI). This result is not confirmed, however, by the experimental data of cyanine "5". Here the first shift is smaller than the two next, which are approximately the same. (These data are not yet very certain in view of the fact that location of the absorption maxima is not clear — see Nikitine (1951)). On the other hand the height of the resonance barrier suggested by the experimental data (as well as by calculations) is extremely large (the greatest shift of absorption maxima toward long waves). Then the probability of "leakage" of electrons across the barriers at the ends of the box outside the nuclei approaches that across the resonance barrier. This suggests another treatment of the problem.

TABLE VI

The stabilization energies and location of absorption maxima for canines in a barrier model

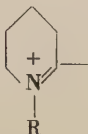
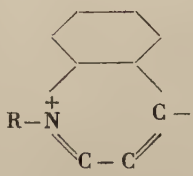
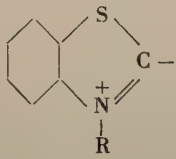
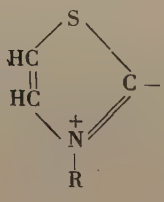
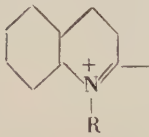
Nr	The nucleus	N	Stabilization energy (eV)	λ (Å) teor.	λ (Å) exp.
"1"		8	13.27	6280	5630
"2"		10	4.91	5640	5910
	"	12	4.91	6955	7050
	"	14	4.91	8260	8130
	"	16	4.91	9565	9300
"3"		6	9.82	3665	4230
	"	8	9.82	5125	5575
	"	10	9.82	6440	6500
	"	12	9.82	7705	7600
	"	14	9.82	8970	8700
	"	16	9.82	10220	9950
"4"		8	12.52	5985	5425

TABLE VI (continued)

Nr	The nucleus	N	Stabilization energy (V).	λ' (Å) teor	λ (Å) exp.
"5"		6	11.78	4300	5235
	..	8	11.78	5725	6040
	..	10	11.78	6925	7080
	..	12	11.78	8265	8200

Only one of the extreme structures of the nuclei is shown in the table. For the experimental data see Brooker (1942), Brober etc. (1941), Kuhn (1948) Nikitine (1951).

For large N the theoretical results of the barrier model are quite close to those of the non-barrier model. It is easily seen, however, that the barrier model as a whole approaches considerably better the experimental data than the simple one (see Table II).

In Sec. 2 we indicated the analogy which existed between the barrier in a free-electron model and the resonance formalism. Let us now compare with the results of this work Moffitt's (1950) conclusions from his detailed analysis of electronic states of the symmetrical cyanines on the basis of the resonance theory. We see that in many points the effect of intermediate structures on the spectrum is similar to the effect which is caused by a barrier in the free-electron scheme.

Thus we have:

1) Moffitt: "the excited structures have greater effect on the ground state than on the excited states".

The present theory: "the introduction of the barrier causes a greater change in a ground state than in an excited level. The higher is the level, the smaller is the effect of the barrier".

2) Moffitt: "with an increase in the chain length the energy difference of the state corresponding to the extreme structures and the state obtained from the interaction of all the structures increases".

The present theory: "with an increase in the chain length at a fixed barrier height, the energy difference of the state of a non-barrier problem and the same state in the presence of the barrier increases".

3) Moffitt: "with the increase of the chain length the mutual stabilization of the ground state and the excited state decreases. The vinylene shift is, therefore, non-convergent".

The present theory: "with the increase in the chain length (and consequently with the increase in n) at a fixed barrier height the situation becomes similar to that in the non-barrier free-electron model. This is known to cause a constant vinylene shift."

КРАТКОЕ СОДЕРЖАНИЕ

С. Олшевский *Теория абсорбционных спектров некоторых линейно сопряженных систем*

В настоящей работе применен метод антисимметрических молекулярных орбит свободного электрона к вычислению абсорбционных максимум в спектре симметрических цианинов и полиенов. Результаты продискутированы в сравнении с эмпирическими данными и результатами простой модели свободного электрона. Потом показано, что можно получить значительное улучшение теоретических ценностей для цианинов посредством введения в модель свободного электрона резонансного барьера потенциала. Высота барьера определена энергией стабилизации ядер цианина. Подробно продискутировано влияние на энергетическое состояние молекулы, вызванное введением в простую модель свободного электрона потенциального барьера. Указано на согласованность полученных этим путем результатов с выводами теории резонанса.

REFERENCES

- Basu S., J. Chem. Phys. **23**, 1625 (1954a); **22**, 1270 (1954b).
 Brooker L.G.S., Rev. Mod. Phys. **14**, 275 (1942); Frontiers in Chemistry, Vol. 3, New York 1945 Interscience Publishers Inc.
 Ecker, White, Keyes, Smyth, Oesper, J. Am. Chem. Soc. **63**, 3192 (1941).
 Kuhn H., Helv. Chim. Acta **31**, 1441 (1948); **34** 2371 (1951); J. Chem. Phys. **17**, 1198 (1949).
 Fukui, Nagata, Yonezawa, J. Chem. Phys. **21**, 186 (1953).
 Moffitt W. E., Proc. Phys. Soc. (London) Sec. A, **63**, 700 (1950).
 Naylor P., Witing M. C., J. Chem. Phys. **22**, 2091 (1954).
 Nikitine S., J. chim. phys. **48**, 37 (1951).
 Olszewski S., Acta phys. Polon. **14**, 419 (1955), *ibid* (to be published)
 Platt J. R., J. Chem. Phys. **22**, 1448 (1954).
 Roothaan C.C.J., Rev. Mod. Phys. **23**, 69 (1951).
 Scherr C. W., J. Chem. Phys. **21**, 1582 (1953).
 Simpson W. T., J. Chem. Phys. **16**, 1124 (1948).

LETTERS TO THE EDITOR

DOMAIN FORMATION IN FERROELECTRICS

BY ARKADIUSZ JAŚKIEWICZ

Physical Laboratory B. Bierut University, Wrocław

(Received September 22, 1956)

The dielectric properties of ferroelectrics depend on the domain structure and domain dynamics. The formation and movement of domains have been studied by various authors: Forsbergh (1949), Merz (1954), Little (1955). Our main aim is to find how the formation of the c domain of BaTiO₃ depends on the electric field.

The nucleation theory supposes that local fluctuations in free energy produce small, ordered regions called embryos. The free energy has to overcome various energy expenditures. Owing to the increase in the surface energy and other expenditures associated with the formation of an embryo, most embryos are unstable and vanish. Embryos greater than a critical size are stable and grow with a decrease in the free energy of the substance. This implies that a critical size can be reached only by the kinetic process of statistical fluctuations see e.g. R. Smoluchowski 1951.

The change in free energy ΔF associated with the formation of a nucleus in ferroelectrics can be represented by the expression

$$\Delta F = \Delta F_w + \Delta F_o + \Delta F_i + \Delta F_d \quad (1)$$

The first term represents the surface energy (wall energy); the second, the energy volume energy (dependent on the parameter of order); the third, the energy of interaction between the polarized dielectric and the field; the last term, the depolarization energy.

If we take the spontaneous polarization as the parameter of order, then for the order energy we obtain the expression

$$\Delta F_o = -\alpha P_s^2 V \quad (2)$$

where α depends on the shape of the embryo, V is the volume, and P_s is the spontaneous polarization.

Formula (1) giving the free energy change in the absence of a field, can be written in the form:

$$\Delta F = \sigma S - \alpha P_s^2 V + \beta P_s^2 V \quad (3)$$

where σ is the wall energy/cm², S the wall area, β the depolarizing factor.

If a phase transformation is to occur, α must be greater than β , because the free energy of the new phase is lower than that of the metastable phase.

We assume that the embryo has a cone-like shape with length l and radius r in accordance with optical observations Hippel (1950), Merz (1954). Using (3) we have

$$\Delta F = \pi\sigma(r^2 + rl) - (\alpha - \beta) \frac{1}{3} \pi r^2 l P_s^2 \quad (4)$$

if we write

$$\alpha - \beta = \gamma \quad (5)$$

we get

$$\Delta F = \pi\sigma(r^2 + rl) - \gamma \frac{1}{3} \pi r^2 l P_s^2 \quad (6)$$

We obtain the critical dimensions from the maximum of (6):

$$r_c = \frac{3\sigma}{\gamma P_s^2}, \quad l_c = \frac{6\sigma}{\gamma P_s^2} \quad (7)$$

We see that there will be only large domains about the transition point. The nucleation rate of domains we can get from well known expression for the rate of nucleation.

Now, we take under consideration a ferroelectric crystal in an external field E . The free energy change of nucleus formation is given by 1, thus, from (6) we can write:

$$\Delta F^P = \pi\sigma(r^2 + rl) - \frac{1}{3} \gamma \pi r^2 l P_s^2 - EP_s V \quad (8a)$$

$$\Delta F^a = \pi\sigma(r^2 + rl) - \frac{1}{3} \gamma \pi r^2 l P_s^2 + EP_s V \quad (8b)$$

where $-EP_s V$ is the electrostatic energy of a nucleus with parallel polarization, $EP_s V$ the energy of a nucleus with antiparallel polarization.

If we repeat our procedure we obtain the critical dimensions:

a) domains with parallel polarization

$$r_c^P = \frac{3\sigma}{P_s(\gamma P_s + E)}; \quad l_c^P = \frac{6\sigma}{P_s(\gamma P_s + E)} \quad (9a)$$

b) domains with antiparallel polarization

$$r_c^a = \frac{3\sigma}{P_s(\gamma P_s - E)}; \quad l_c^a = \frac{6\sigma}{P_s(\gamma P_s - E)} \quad (9b)$$

We see that equations (9a) and (9b) explain the behaviour of domains in the electric field. In the case of a nucleus with antiparallel polarization, the critical dimensions are greater than in the case of parallel polarization. All the antiparallel nuclei which have dimensions less than the critical one are unstable and vanish. The destroyed antiparallel domains have a greater probability to nucleate parallelly than antiparallely. The probability may be obtained from the nucleation rate. The last equations are well suited for a study of domain formation in connection with experiments such as those of Piekara and Pająk 1953 and of Badian and Jaśkiewicz.

The author wishes to express his thanks to Dr. Z. Galasiewicz and Professor B. Makiej for helpful discussions.

REFERENCES

- Badian, L., and Jaśkiewicz, A., To be published.
Forsbergh, P. W., *Phys. Rev.* **76**, 1187 (1949).
Hippel, A., *Rev. Mod. Phys.* **22**, 221 (1950).
Little, E. A. *Phys. Rev.* **98**, 978 (1955).
Merz, W. J. *Phys. Rev.* **95**, 690 (1954).
Piekara, A and Pająk, Z., *Acta phys. Polon.* **12**, 170 (1953).
Smoluchowski, R., *Phase Transformations in Solids*, New York, London (1951).

NOTE ON THE $^{12}\text{C}(\text{d},\text{p})^{13}\text{C}$ REACTION NEAR THE 4 MeV RESONANCE

J. DĄBROWSKI AND B. TULCZYJEW

Institute of Theoretical Physics, University of Warsaw, Warsaw

(Received February 11, 1957)

Recently several observations have been performed for the stripping reactions in the vicinity of an isolated resonance of the compound nucleus. A calculation scheme for these reactions is outlined in the present note in its most simple version for the case of the $^{12}\text{C}(\text{d},\text{p})^{13}\text{C}$ reaction near the 4 MeV resonance investigated experimentally by Bonner et al. (1956a).¹ We adopt their values: $J = 3^+$ and the width $\Gamma_{lab} = 0.15$ MeV of the excited level in ^{14}N .

The formula for the cross-section for the reactions of the type considered here has been given by Thomas (1955) (cf. also Dąbrowski 1956²). To perform the numerical calculations on the basis of this formula we make the following simplifying assumptions:

- (1) The Coulomb effect is not taken into account.
- (2) The potential scattering of deuteron and proton is neglected.
- (3) Following the suggestion of Bonner et al. (1956a) we assume the values $l_d = 2$, $j_p = 0$ (channel spin) as responsible for the compound nucleus mechanism. Thus we neglect the contribution of $l_d = 4$ and $j_p = 1$.
- (4) The neutron is captured with $l_n = 1$, this being in agreement with the shell model.

Owing to these assumptions we may write the differential cross-section for the $^{12}\text{C}(\text{d},\text{p})^{13}\text{C}$ reaction in the vicinity of the 4 MeV resonance in the form

$$d\sigma/d\Omega = (\gamma^2/a) [(d\sigma/d\Omega)_S + \lambda(d\sigma/d\Omega)_I + \lambda^2(d\sigma/d\Omega)_C],$$

where γ^2 is the reduced width of the neutron bound state in ^{13}C ; $\lambda = \gamma_p\gamma_d/\gamma$; γ_p^2 , γ_d^2 are the corresponding partial widths of the virtual level in ^{14}N ; $\gamma^2 a^{-1}(d\sigma/d\Omega)_S$ is the usual stripping cross-section; $\gamma^2 a^{-1}\lambda(d\sigma/d\Omega)_I$ is the interference cross-section (constructed as in Eq. (17) of the paper by Thomas (1955); $\gamma^2 a^{-1}\lambda^2(d\sigma/d\Omega)_C$ is the Breit-

¹ It is to be noticed that in the paper by Bonner et al. (1956a) the energies given in Fig. 3 should be reduced by about 0.08 MeV and that the solid curve in Fig. 11 is not correct (it should have no minima for 0° and 180°).

² The symbols used here are the same as in that paper.

Wigner — one level — cross-section (without the Coulomb and potential scattering phase shifts).

The calculations were performed with the help of Butler's theory and Born's theory (transparent nucleus). It was found that the best value of the (stripping) radius of the ^{12}C nucleus is $a \approx 6 \times 10^{-13}$ cm (for the Born theory still a larger value of a should be used).

Now let us discuss the value of the resonance energy. Owing to the observation of Bonner et al. (1956a) the deuteron resonance energy $E_r^{obs} = 4.004$ MeV, and the occurrence of the resonance is most distinct in the first stripping maximum. The interference cross-section causes that the true resonance energy E_r differs from E_r^{obs} . For small λ we get the maximum value of $d\sigma/d\Omega$ in the first stripping maximum for $E_r^{obs} \approx E_r \pm \Gamma/2$. In our calculation the plus sign is used ($E_r = 3.93$ MeV) since in the reactions $^{10}\text{B}(\alpha, n)^{13}\text{N}$ and $^{10}\text{B}(\alpha, p\gamma)^{13}\text{C}$ (Bonner et al. 1956b) the corresponding resonance energy was observed to be equal to 3.98 MeV, and it is hoped that in the case of both these reactions the role of the direct interaction is smaller and the observed resonance energy is closer to the true resonance energy than in the case of the (d,p) reaction.

The quantities λ , γ were determined by fitting the values of the theoretical cross-section at the first stripping maximum to the experimental values at $E_d = 3.84$ and 4.01 MeV. We found

$$\gamma^2 = \begin{cases} 0.87 \times 10^{-13} \text{ MeV cm,} \\ 0.22 \times 10^{-13} \text{ MeV cm,} \end{cases} \quad \gamma_p \gamma_d = \begin{cases} 0.0099 \text{ MeV, for the Butler theory,} \\ 0.0012 \text{ MeV, for the Born theory.} \end{cases}$$

The corresponding contribution of $(d\sigma/d\Omega)_C$ to $d\sigma/d\Omega$ at the first maximum is about 3% for both the energies. It should be noticed that the dimensionless reduced width $\theta^2 = 2M^* a \gamma^2 / \hbar^2$ in the case of the Born calculation (0.058) is in rather close agreement with the value 0.07 given by Thomas (1952).

The corresponding angular distributions are given in Fig. 1. It is hardly possible to decide whether the Butler or the Born angular distribution is better. We see from Fig. 1. that the compound nucleus mechanism is not strong enough to give the observed back scattering maximum. This is still more evident from Fig. 2. for $E_d = 4.12$ MeV. On the other hand, for this energy the compound nucleus mechanism is too large, since the calculated value of the cross-section at the first maximum is larger than the experimental value.

Assumptions (1), (2) are very crude. One may argue that the influence of the Coulomb effect and potential scattering has been investigated by several authors (Horowitz and Messiah 1953, Yoccoz 1954, Grant 1945—1955, Tobocman 1954, Tobocman and Kalos 1955). But the neglect of these effects in the interference cross-section is a far reaching simplification, since it follows from it that $(d\sigma/d\Omega)_I \sim (E_d - E_r)$. Of course, the term proportional to $(E_d - E_r)$ will remain in a more exact formula also and will cause a shift of the observed resonance.

Assumption (3) is somewhat arbitrary. It is to be noticed, however, that an introduction of other angular momenta into the calculation would introduce new parameters (the corresponding partial widths) automatically³. This would destroy the simplicity of the present treatment.

The results of Bonner et al. (1956a) suggest the vanishing of $(d\sigma/d\Omega)_I$ near the first stripping maximum (cf. their Fig. 7). This is not the case in our calculation. Connected with this is our rather unrealistic assumption $E_r = 3.93$ MeV. All these facts show that our simplifying assumptions are going too far. The spin value 3^+ of the resonance level is rather arbitrary also. Nevertheless it is hoped that one may conclude already on the basis of this calculation that the compound nucleus formation is not a sufficient reason for the back scattering maximum even in the vicinity of the reso-

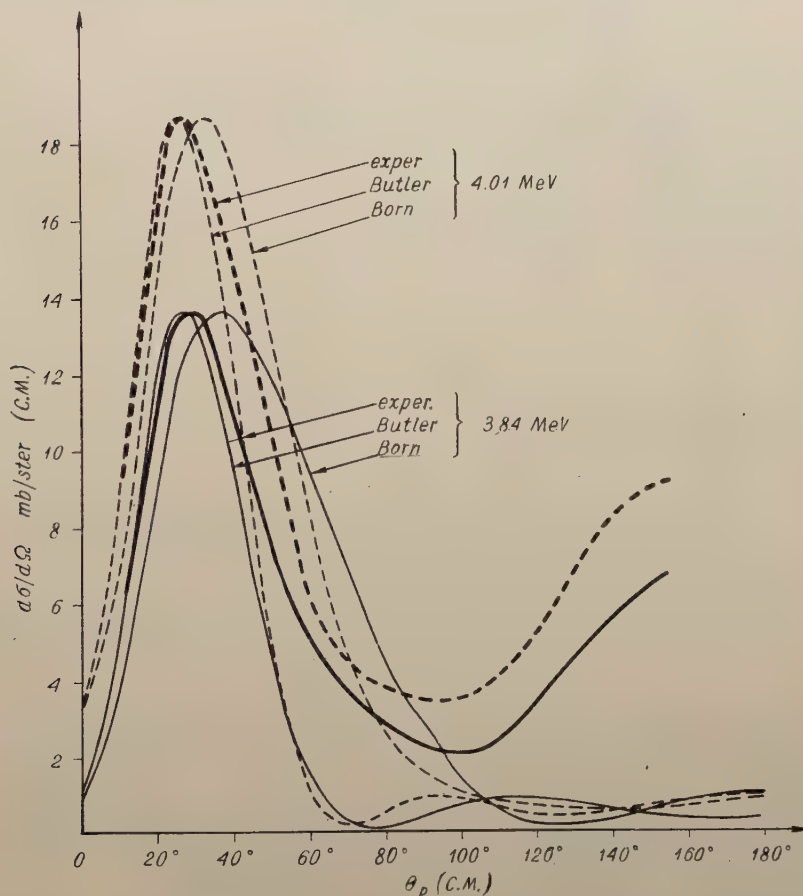


Fig. 1. Angular distribution of protons from $^{12}\text{C}(\text{d},\text{p})^{13}\text{C}$ for the deuteron energy $E_d = 3.84$ and 4.01 MeV

³ The situation would be simpler in the case of spinless particles, where the number of the possible combinations of the angular momenta becomes smaller. For such cases the scheme presented here would be more useful.

nance. Further it is hoped that the scheme presented may be of some value for other similar calculations.

We are indebted to Dr. J. B. Marion for explanatory remarks concerning his papers (with Bonner et al.) and the problem discussed in this note.

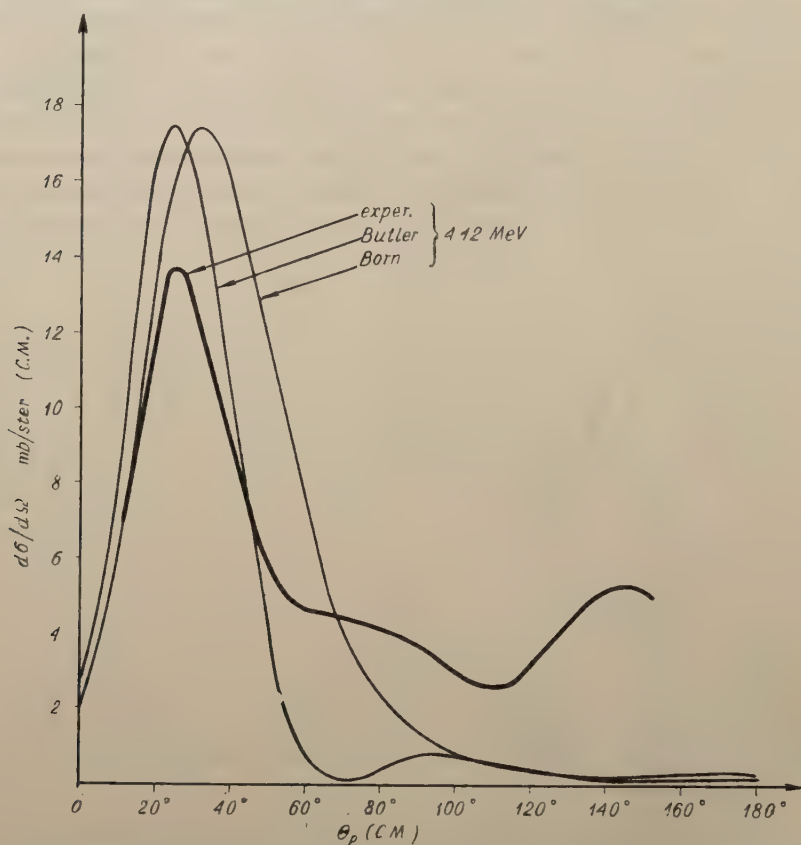


Fig. 2. Angular distribution of protons from $^{12}\text{C}(\text{d},\text{p})^{13}\text{C}$ for the deuteron energy $E_d = 4.12$ MeV

REFERENCES

- Bonner T. W., Eisinger J. T., Kraus A. A. Jr. and Marion J. B., *Phys. Rev.*, **101**, 209 (1956a).
 Bonner T. W., Kraus A. A. Jr., Marion J. B., and Schiffer J. P., *Phys. Rev.*, **102**, 1348 (1956b).
 Dąbrowski J., *Acta phys. Polon.*, (1956).
 Grant I. P., *Proc. Phys. Soc. A.*, **67**, 981 (1954); *ibid.*, **68**, 244 (1955).
 Horowitz J., and Messiah A. M. L., *J. Phys. Rad.*, **14**, 695 (1953).
 Thomas R. G., *Phys. Rev.*, **88**, 1109 (1952); *ibid.*, **100**, 25 (1955).
 Tobocman W., *Phys. Rev.*, **94**, 1655 (1954).
 Tobocman W., and Kalos M. H., *Phys. Rev.*, **97**, 132 (1955).

APPENDIX IV

TO THE PAPER BY J. SAWICKI

"Neutron Polarization and Angular Distribution in (p, n) Reactions",Acta phys. Polon. **1—2**, 95 (1957)

(Received April 13, 1957)

Derivation of the Final State of the Outgoing Neutron Eq. (3) represents the "equivalent" outgoing state of the neutron scattering in a spin-orbit potential its radial partial waves being proportional to $J_L(k_n r) - \beta(L, J) h_L^{(1)}(k_n r)$. Eq. (3) results as a generalization of Tobočan's Eq. (2) for spinless particles (1954, Eq. (19)). In analogy to Tobočan's Eq. (2) for spinless particles let us introduce the spin-dependent function:

$$f_{\mu_f}(\vec{r}_n, \sigma_n) = \int d\sigma_p d\vec{r}_p d\xi \Psi_f^*[\mu_f](\xi, \vec{r}_p, \sigma_p) \Psi \quad (1')$$

Ψ being the exact three-body function of the problem.

In analogy to Tobočan's Eq. (11) we may write the integral equation for (1'):

$$f_{\mu_f}(\vec{r}_n, \sigma_n) = - \frac{M}{2\pi\hbar^2} \int d\vec{r}_n' \frac{e^{ik_n|\vec{r}_n - \vec{r}_n'|}}{|\vec{r}_n - \vec{r}_n'|} V_{Fn} f_{\mu_f}(\vec{r}_n', \sigma_n) \quad (2')$$

V_{Fn} being an integral operator. We can expand $f_{\mu_f}(\vec{r}_n, \sigma_n)$ in total angular momentum eigenfunctions:

$$f_{\mu_f}(\vec{r}_n, \sigma_n) = \sum_{L, M_L} \sum_{J, M_J} f_{LJ}^{\mu_f}(r_n) \mathcal{Y}_{JL \frac{1}{2}}^{M_J}(\Omega_n, \sigma_n) \quad (3')$$

where:

$$\mathcal{Y}_{JL \frac{1}{2}}^{M_J}(\Omega_n, \sigma_n) = \sum_{M'_L, \mu_n} \left(L, \frac{1}{2}; M'_L, \mu_n | J, M_J \right) Y_{LM'_L}(\Omega_n) \chi_{\mu_n}(\sigma_n) \quad (4')$$

Introducing the expansion (3') into Eq. (2') gives as a final result:

$$f_{LJ}^{\mu_f}(r_n) = \frac{2k_n M}{i\hbar^2} \int d\vec{r}_n' d\sigma_n J_L(k_n r_n) h_L^{(1)}(k_n r_n) \mathcal{Y}_{JL \frac{1}{2}}^{M_J}(\Omega_n', \sigma_n) V_{Fn} f_{\mu_f}(\vec{r}_n', \sigma_n) \quad (5')$$

where:

$$r_{<} = \begin{cases} r'_n & \text{for } r'_n < r_n \\ r_n & \text{for } r'_n > r_n \end{cases} \text{ and } r_{>} = \begin{cases} r'_n & \text{for } r'_n > r_n \\ r_n & \text{for } r'_n < r_n \end{cases}$$

One can now easily eliminate the interior region of the nucleus and to derive the expression for the distortion coefficient $\beta^{\mu_f}(L, J)$ in complete analogy to Tobocman's expression for β_{fLM}^* (Eq. (17)):

$$\beta^{\mu_f}(L, J) = \left[\frac{\frac{\partial}{\partial r} J_L(k_n r) - J_L(k_n r) \frac{\partial}{\partial r} \ln f_{LJ}^{\mu_f}(r)}{\frac{\partial}{\partial r} h_L^{(1)}(k_n r) - h_L^{(1)}(k_n r) \frac{\partial}{\partial r} \ln f_{LJ}^{\mu_f}(r)} \right]_{r=R_0} \quad (6')$$

If now, we indentify $\left[\frac{\partial}{\partial r} \ln f_{LJ}^{\mu_f}(r) \right]_{R_0}$ with the logarithmic derivative appropriate for free neutrons incident on the residual nucleus, the μ_f -dependence is neglected and we obtain $\beta(L, J) = \frac{1 - \bar{\eta}_{LJ}}{2}$, $\bar{\eta}_{LJ}$ being the average reflection coefficients in analogy to Eq. (II,5) — (II-7).

Volumen XVI (1957) — Fasciculus 3

L. Infeld, The Equations of Motion in General Relativity Theory and the Action Principle	177
S. Olszewski, On the Free-Electron Theory of Absorption Spectra of Some Linear Conjugated Systems	211

Letters to the Editor:

A. Jaśkiewicz, Domain Formation in Ferroelectrics	227
J. Dąbrowski and B. Tulczyjew, Note on the $^{12}\text{C}(\text{d,p})^{13}\text{C}$ Reaction Near the 4 MeV Resonance	231